

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:		(11) International Publication Number:	WO 99/42120
A61K 38/00, C07K 5/00, 7/00, 16/00, 17/00, C07H 21/04, C12N 15/00, 15/09, 15/63, 15/70, 15/74, C12P 21/06, C12Q 1/68	A1	(43) International Publication Date:	26 August 1999 (26.08.99)
(21) International Application Number: PCT/US9 (22) International Filing Date: 18 February 1999 (1		CH, CY, DE, DK, ES, FI, FR, (European patent (AT, BE, GB, GR, IE, IT, LU, MC,
(30) Priority Data: 09/027,337 20 February 1998 (20.02.98)	U:	Published With international search report.	· · · · · · · · · · · · · · · · · · ·
(71) Applicant: THE BOARD OF TRUSTEES OF THE USITY OF ARKANSAS [US/US]; 2404 North University enue, Little Rock, AR 72207-3608 (US).		•	
(72) Inventors: O'BRIEN, Timothy, J.; 2625 Grist Mill Roa Rock, AR 72227 (US). TANIMOTO, Hirotoshi; Ap 804, 701 Green Mountain Drive, Little Rock, AR (US).	partmen	t l	
74) Agent: ADLER, Benjamin, A.; McGregor & Adle Candle Lane, Houston, TX 77071 (US).	r, 8011		

(54) Title: TADG-15: AN EXTRACELLULAR SERINE PROTEASE OVEREXPRESSED IN BREAST AND OVARIAN CARCINOMAS

(57) Abstract

The present invention provides a DNA encoding a TADG-15 protein selected from the group consisting of: (a) isolated DNA which encodes a TADG-15 protein; (b) isolated DNA which hybridizes to isolated DNA of (a) above and which encodes a TADG-15 protein; and (c) isolated DNA differing from the isolated DNAs of (a) and (b) above in codon sequence due to the degeneracy of the genetic code, and which encodes a TADG-15 protein. Also provided is a vector capable of expressing the DNA of the present invention adapted for expression in a recombinant cell and regulatory elements necessary for expression of the DNA in the cell.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	£U	Luxembourg	SN	Senegal
ΑÜ	Australia	GA	Gabon	ŁV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB.	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgiuni	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	ĮΕ	Ireland	MN	Mongolia	ÜA .	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan-
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Кепуа	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	КG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL ·	Poland	_	
CN	China	KR	Republic of Korea	PT	Portugal	•	
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DΚ	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		·

TADG-15: AN EXTRACELLULAR SERINE PROTEASE OVEREXPRESSED IN BREAST AND OVARIAN CARCINOMAS

10

BACKGROUND OF THE INVENTION

15 Field of the Invention

The present invention relates generally to the fields of cellular biology and the diagnosis of neoplastic disease. More specifically, the present invention relates to an extracellular serine protease termed Tumor Antigen Derived Gene-15 (TADG-15), which is overexpressed in breast and ovarian carcinomas.

Description of the Related Art

Extracellular proteases have been directly associated with tumor growth, shedding of tumor cells and invasion of target organs.

25 Individual classes of proteases are involved in, but not limited to (1) the digestion of stroma surrounding the initial tumor area, (2) the digestion of the cellular adhesion molecules to allow dissociation of tumor cells; and (3) the invasion of the basement membrane for

metastatic growth and the activation of both tumor growth factors and angiogenic factors.

The prior art is deficient in the lack of effective means of screening to identify proteases overexpressed in carcinoma. The present invention fulfills this longstanding need and desire in the art.

SUMMARY OF THE INVENTION

The present invention discloses a screening program to identify proteases overexpressed in carcinoma by examining PCR products amplified using differential display in early stage tumors, metastatic tumors compared to that of normal tissues.

15

20

25

In one embodiment of the present invention, there is provided a DNA encoding a TADG-15 protein selected from the group consisting of: (a) isolated DNA which encodes a TADG-15 protein; (b) isolated DNA which hybridizes to isolated DNA of (a) above and which encodes a TADG-15 protein; and (c) isolated DNA differing from the isolated DNAs of (a) and (b) above in codon sequence due to the degeneracy of the genetic code, and which encodes a TADG-15 protein.

In another embodiment of the present invention, there is provided a vector capable of expressing the DNA of the present invention adapted for expression in a recombinant cell and regulatory elements necessary for expression of the DNA in the cell.

In yet another embodiment of the present invention, there is provided a host cell transfected with the vector of the present invention, the vector expressing a TADG-15 protein.

In still yet another embodiment of the present invention, there is provided a method of detecting expression of a TADG-15 mRNA, comprising the steps of: (a) contacting mRNA obtained from the cell with the labeled hybridization probe; and (b) detecting hybridization of the probe with the mRNA.

Other and further aspects, features, and advantages of the present invention will be apparent from the following description of the presently preferred embodiments of the invention given for the purpose of disclosure.

10

15

20

BRIEF DESCRIPTION OF THE DRAWINGS

So that the matter in which the above-recited features, advantages and objects of the invention, as well as others which will become clear, are attained and can be understood in detail, more particular descriptions of the invention briefly summarized above may be had by reference to certain embodiments thereof which are illustrated in the appended drawings. These drawings form a part of the specification. It is to be noted, however, that the appended drawings illustrate preferred embodiments of the invention and therefore are not to be considered limiting in their scope.

Figure 1 shows a comparison of PCR products derived from normal and breast carcinoma cDNA as shown by staining in an agarose gel.

Figure 2 shows a comparison of the serine protease catalytic domain of TADG-15 (SEQ ID No: 14) with hepsin (Heps, SEQ ID No: 3), (Scce, SEQ ID No: 4), trypsin (Try, SEQ ID No: 5), chymotrypsin (Chymb, SEQ ID No: 6), factor 7 (Fac7, SEQ ID No:

7) and tissue plasminogen activator (Tpa, SEQ ID No: 8). The asterisks indicate conserved amino acids of catalytic triad.

Figure 3 shows quantitative PCR analysis of TADG-15 expression.

Figure 4 shows the ratio of TADG-15 expression to expression of β -tubulin in normal tissues, low malignant potential tumors (LMP) and carcinomas.

Figure 5 shows the TADG-15 expression in tumor cell lines derived from both ovarian and breast carcinoma tissues.

Figure 6 shows the overexpression of TADG-15 in other tumor tissues.

Figure 7 shows the Northern blots of TADG-15 expression in ovarian carcinomas, fetal and normal adult tissues.

Figure 8 shows a diagram of the TADG-15 transcript and the clones with the origin of their derivation.

Figure 9 shows nucleotide sequence of the TADG-15 cDNA (SEQ ID No: 1) and amino acid sequence of the TADG-15 protein (SEQ ID No: 2)

Figure 10 shows the amino acid sequence of the TADG-20 15 protease including functional sites and domains.

Figure 11 shows a structure diagram of the TADG-15 protein including functional domains.

Figure 12 shows a nucleotide sequence comparison between TADG-15 and human SNC-19 (GeneBank accession #U20428).

25

15

5

DETAILED DESCRIPTION OF THE INVENTION

As used herein, the term "cDNA" shall refer to the DNA copy of the mRNA transcript of a gene.

As used herein, the term "derived amino acid sequence" shall mean the amino acid sequence determined by reading the triplet sequence of nucleotide bases in the cDNA.

As used herein the term "screening a library" shall refer to the process of using a labeled probe to check whether, under the appropriate conditions, there is a sequence complementary to the probe present in a particular DNA library. In addition, "screening a library" could be performed by PCR.

10

15

20

25

As used herein, the term "PCR" refers to the polymerase chain reaction that is the subject of U.S. Patent Nos. 4,683,195 and 4,683,202 to Mullis, as well as other improvements now known in the art.

The TADG-15 cDNA is 3147 base pairs long (SEQ ID No:1) and encoding for a 855 amino acid protein (SEQ ID No:2). The availability of the TADG-15 gene opens the way for a number studies that can lead to various applications. For example, the TADG-15 gene can be used as a diagnostic or therapeutic target in ovarian carcinoma and other carcinomas including breast, prostate, lung and colon.

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Maniatis, Fritsch & Sambrook, "Molecular Cloning: A Laboratory Manual (1982); "DNA Cloning: A Practical Approach," Volumes I and II (D.N. Glover ed. 1985); "Oligonucleotide Synthesis" (M.J. Gait ed. 1984); "Nucleic Acid Hybridization" [B.D. Hames & S.J. Higgins eds. (1985)];

"Transcription and Translation" [B.D. Hames & S.J. Higgins eds. (1984)];
"Animal Cell Culture" [R.Î. Freshney, ed. (1986)]; "Immobilized Cells
And Enzymes" [IRL Press, (1986)]; B. Perbal, "A Practical Guide To
Molecular Cloning" (1984).

Therefore, if appearing herein, the following terms shall have the definitions set out below.

The amino acid described herein are preferred to be in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the desired functional property of immunoglobulin-binding is retained by the polypeptide. NH2 refers to the free amino group present at the amino terminus of a polypeptide. COOH refers to the free carboxy group present at the carboxy terminus of a polypeptide. In keeping with standard polypeptide nomenclature, *J Biol. Chem.*, 243:3552-59 (1969), abbreviations for amino acid residues are shown in the following Table of Correspondence:

10

15

TABLE OF CORRESPONDENCE

	SYMBOL			AMINO ACID
	<u>l-Letter</u>		3-Letter	
5	Y		Tyr	tyrosine
	G	· ·	Gly	glycine
	F		Phe	Phenylalanine
•	M		Met	methionine
	Α	•	Ala	alanine
10	S		Ser	serine
	I		Ile	isoleucine
	L		Leu	leucine
	T		Thr	threonine
	V		. Val	valine
15	P		Pro	proline
	K		Lys	lysine
	H	•	His	histidine
	Q	•	Gln	glutamine
	E		Glu	glutamic acid
20	W		Trp	tryptophan
•	R		Arg	arginine
	D		Asp	aspartic acid
	N	•	Asn	asparagine
	С	•	Cys	cysteine
		•		

25

30

It should be noted that all amino-acid residue sequences are represented herein by formulae whose left and right orientation is in the conventional direction of amino-terminus to carboxy-terminus. Furthermore, it should be noted that a dash at the beginning or end of an amino acid residue sequence indicates a peptide bond to a further sequence of one or more amino-acid residues. The above Table is presented to correlate the three-letter and one-letter notations which may appear alternately herein.

A "replicon" is any genetic element (e.g., plasmid, chromosome, virus) that functions as an autonomous unit of DNA replication in vivo; i.e., capable of replication under its own

control.

10

15

20

25

A "vector" is a replicon, such as plasmid, phage or cosmid, to which another DNA segment may be attached so as to bring about the replication of the attached segment.

A "DNA molecule" refers to the polymeric form of deoxyribonucleotides (adenine, guanine, thymine, or cytosine) in its either single stranded form, or a double-stranded helix. This term refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary forms. Thus, this term includes double-stranded DNA found, *inter alia*, in linear DNA molecules (e.g., restriction fragments), viruses, plasmids, and chromosomes. In discussing the structure herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the nontranscribed strand of DNA (i.e., the strand having a sequence homologous to the mRNA).

An "origin of replication" refers to those DNA sequences that participate in DNA synthesis.

A DNA "coding sequence" is a double-stranded DNA sequence which is transcribed and translated into a polypeptide in vivo when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxyl) terminus. A coding sequence can include, but is not limited to, prokaryotic sequences, cDNA from eukaryotic mRNA, genomic DNA sequences from eukaryotic (e.g., mammalian) DNA, and even synthetic DNA sequences. A polyadenylation signal and transcription termination sequence will usually be located 3' to the coding sequence.

Transcriptional and translational control sequences are DNA regulatory sequences, such as promoters, enhancers, polyadenylation signals, terminators, and the like, that provide for the expression of a coding sequence in a host cell.

5

10

15

20

A "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements initiate transcription at levels detectable necessary to above background. Within the promoter sequence will be found a transcription initiation site, as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. Eukaryotic promoters often, but not always, contain "TATA" boxes and "CAT" boxes. Prokaryotic promoters contain Shine-Dalgarno sequences in addition to the -10 and -35 consensus sequences.

An "expression control sequence" is a DNA sequence that controls and regulates the transcription and translation of another DNA sequence. A coding sequence is "under the control" of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then translated into the protein encoded by the coding sequence.

A "signal sequence" can be included near the coding sequence. This sequence encodes a signal peptide, N-terminal to the polypeptide, that communicates to the host cell to direct the polypeptide to the cell surface or secrete the polypeptide into the media, and this signal peptide is clipped off by the host cell

before the protein leaves the cell. Signal sequences can be found associated with a variety of proteins native to prokaryotes and eukaryotes.

The term "oligonucleotide", as used herein in referring to the probe of the present invention, is defined as a molecule comprised of two or more ribonucleotides, preferably more than three. Its exact size will depend upon many factors which, in turn, depend upon the ultimate function and use of the oligonucleotide.

10

15

20

"primer" as used herein refers The term oligonucleotide, whether occurring naturally as in a purified restriction digest or produced synthetically, which is capable of acting as a point of initiation of synthesis when placed under conditions in synthesis of a primer extension product, which which complementary to a nucleic acid strand, is induced, i.e., in the presence of nucleotides and an inducing agent such as a DNA polymerase and at a suitable temperature and pH. The primer may be either single-stranded or double-stranded and must be sufficiently long to prime the synthesis of the desired extension product in the presence of the inducing agent. The exact length of the primer will depend upon many factors, including temperature, source of primer and use the method. For example, for diagnostic applications, depending on the complexity of the target sequence, the oligonucleotide primer typically contains 15-25 or more nucleotides, although it may contain fewer nucleotides.

The primers herein are selected to be "substantially" complementary to different strands of a particular target DNA sequence. This means that the primers must be sufficiently complementary to hybridize with their respective strands.

Therefore, the primer sequence need not reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment may be attached to the 5' end of the primer, with the remainder of the primer sequence being complementary to the strand. Alternatively, non-complementary bases or longer sequences can be interspersed into the primer, provided that the primer sequence has sufficient complementary with the sequence or hybridize therewith and thereby form the template for the synthesis of the extension product.

As used herein, the terms "restriction endonucleases" and "restriction enzymes" refer to enzymes, each of which cut double-stranded DNA at or near a specific nucleotide sequence.

10

20

25

A cell has been "transformed" by exogenous or heterologous DNA when such DNA has been introduced inside the cell. The transforming DNA may or may not be integrated (covalently linked) into the genome of the cell. In prokaryotes, yeast, and mammalian cells for example, the transforming DNA may be maintained on an episomal element such as a plasmid. With respect to eukaryotic cells, a stably transformed cell is one in which the transforming DNA has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transforming DNA. A "clone" is a population of cells derived from a single cell or ancestor by mitosis. A "cell line" is a clone of a primary cell that is capable of stable growth in vitro for many generations.

Two DNA sequences are "substantially homologous"

when at least about 75% (preferably at least about 80%, and most preferably at least about 90% or 95%) of the nucleotides match over the defined length of the DNA sequences. Sequences that are substantially homologous can be identified by comparing the sequences using standard software available in sequence data banks, or in a Southern hybridization experiment under, for example, stringent conditions as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, e.g., Maniatis et al., supra; DNA Cloning, Vols. I & II, supra; Nucleic Acid Hybridization, supra.

10

15

20

25

A "heterologous' region of the DNA construct is an identifiable segment of DNA within a larger DNA molecule that is not found in association with the larger molecule in nature. Thus, when the heterologous region encodes a mammalian gene, the gene will usually be flanked by DNA that does not flank the mammalian genomic DNA in the genome of the source organism. In another example, coding sequence is a construct where the coding sequence itself is not found in nature (e.g., a cDNA where the genomic coding sequence contains introns, or synthetic sequences having codons different than the native gene). Allelic variations or naturally-occurring mutational events do not give rise to a heterologous region of DNA as defined herein.

The labels most commonly employed for these studies are radioactive elements, enzymes, chemicals which fluoresce when exposed to ultraviolet light, and others. A number of fluorescent materials are known and can be utilized as labels. These include, for example, fluorescein, rhodamine, auramine, Texas Red, AMCA blue and Lucifer Yellow. A particular detecting material is anti-rabbit

antibody prepared in goats and conjugated with fluorescein through an isothiocyanate.

Proteins can also be labeled with a radioactive element or with an enzyme. The radioactive label can be detected by any of the currently available counting procedures. The preferred isotope may be selected from ³H, ¹⁴C, ³²P, ³⁵S, ³⁶Cl, ⁵¹Cr, ⁵⁷Co, ⁵⁸Co, ⁵⁹Fe, ⁹⁰Y, ¹²⁵I, ¹³¹I, and ¹⁸⁶Re.

Enzyme labels are likewise useful, and can be detected by any of the presently utilized colorimetric, spectrophotometric, fluorospectrophotometric, amperometric or gasometric techniques. The enzyme is conjugated to the selected particle by reaction with carbodiimides, diisocyanates, molecules such as bridging glutaraldehyde and the like. Many enzymes which can be used in these procedures are known and can be utilized. The preferred are peroxidase, β-glucuronidase, β-D-glucosidase, β-D-galactosidase, urease, glucose oxidase plus peroxidase and alkaline phosphatase. U.S. Patent Nos. 3,654,090, 3,850,752, and 4,016,043 are referred to by way of example for their disclosure of alternate labeling material and methods.

10

15

20

25

A particular assay system developed and utilized in the art is known as a receptor assay. In a receptor assay, the material to be assayed is appropriately labeled and then certain cellular test colonies are inoculated with a quantitiy of both the label after which binding studies are conducted to determine the extent to which the labeled material binds to the cell receptors. In this way, differences in affinity between materials can be ascertained.

An assay useful in the art is known as a "cis/trans" assay.

Briefly, this assay employs two genetic constructs, one of which

is typically a plasmid that continually expresses a particular receptor of interest when transfected into an appropriate cell line, and the second of which is a plasmid that expresses a reporter luciferase, under the control of a receptor/ligand complex. Thus, for example, if it is desired to evaluate a compound as a ligand for a particular receptor, one of the plasmids would be a construct that results in expression of the receptor in the chosen cell line, while the second plasmid would possess a promoter linked to the luciferase gene in which the response element to the particular receptor is inserted. If the compound under test is an agonist for the receptor, the ligand will complex with the receptor, and the resulting complex will bind the response element and initiate transcription of the The resulting chemiluminescence is then measured luciferase gene. obtained photometrically, and dose response curves are compared to those of known ligands. The foregoing protocol is described in detail in U.S. Patent No. 4,981,784.

1.0

15

20

25

As used herein, the term "host" is meant to include not only prokaryotes but also eukaryotes such as yeast, plant and animal cells. A recombinant DNA molecule or gene which encodes a human TADG-15 protein of the present invention can be used to transform a host using any of the techniques commonly known to those of ordinary skill in the art. Especially preferred is the use of a vector containing coding sequences for the gene which encodes a human TADG-15 protein of the present invention for purposes of prokaryote transformation. Prokaryotic hosts may include *E. coli*, *S. tymphimurium*, *Serratia marcescens* and *Bacillus subtilis*. Eukaryotic hosts include yeasts such as *Pichia pastoris*, mammalian cells and insect cells.

In general, expression vectors containing promoter sequences which facilitate the efficient transcription of the inserted DNA fragment are used in connection with the host. The expression vector typically contains an origin of replication, promoter(s), terminator(s), as well as specific genes which are capable of providing phenotypic selection in transformed cells. The transformed hosts can be fermented and cultured according to means known in the art to achieve optimal cell growth.

The invention includes a substantially pure DNA encoding a TADG-15 protein, a strand of which DNA will hybridize at high stringency to a probe containing a sequence of at least 15 consecutive nucleotides of (SEQ ID NO:1). The protein encoded by the DNA of this invention may share at least 80% sequence identity (preferably 85%, more preferably 90%, and most preferably 95%) with the amino acids listed in Figure 10 (SEQ ID NO:2). More preferably, the DNA includes the coding sequence of the nucleotides of Figure 9 (SEQ ID NO:1), or a degenerate variant of such a sequence.

10

20

25

The probe to which the DNA of the invention hybridizes preferably consists of a sequence of at least 20 consecutive nucleotides, more preferably 40 nucleotides, even more preferably 50 nucleotides, and most preferably 100 nucleotides or more (up to 100%) of the coding sequence of the nucleotides listed in Figure 9 (SEQ ID NO:1) or the complement thereof. Such a probe is useful for detecting expression of TADG-15 in a human cell by a method including the steps of (a) contacting mRNA obtained from the cell with the labeled hybridization probe; and (b) detecting hybridization of the probe with the mRNA.

This invention also includes a substantially pure

DNA containing a sequence of at least 15 consecutive nucleotides (preferably 20, more preferably 30, even more preferably 50, and most preferably all) of the region from nucleotides 1 to 3147 of the nucleotides listed in Figure 9 (SEQ ID NO:1).

By "high stringency" is meant DNA hybridization and wash conditions characterized by high temperature and low salt concentration, e.g., wash conditions of 65°C at a salt concentration of approximately 0.1 x SSC, or the functional equivalent thereof. For example, high stringency conditions may include hybridization at about 42°C in the presence of about 50% formamide; a first wash at about 65°C with about 2 x SSC containing 1% SDS; followed by a second wash at about 65°C with about 0.1 x SSC.

10

15

20

25

By "substantially pure DNA" is meant DNA that is not part of a milieu in which the DNA naturally occurs, by virtue of separation (partial or total purification) of some or all of the molecules of that milieu, or by virtue of alteration of sequences that flank the claimed DNA. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote; or which exists as a separate molecule (e.g., a cDNA or a genomic or cDNA fragment produced by polymerase chain reaction (PCR) or restriction endonuclease digestion) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence, e.g., a fusion protein. Also included is a recombinant DNA which includes a portion of the nucleotides listed in Figure 9 (SEQ ID NO:1) which encodes an alternative splice variant of TADG-15.

The DNA may have at least about 70% sequence

identity to the coding sequence of the nucleotides listed in Figure 9 (SEQ ID NO:1), preferably at least 75% (e.g. at least 80%); and most preferably at least 90%. The identity between two sequences is a direct function of the number of matching or identical positions. When a subunit position in both of the two sequences is occupied by the same monomeric subunit, e.g., if a given position is occupied by an adenine in each of two DNA molecules, then they are identical at For example, if 7 positions in a sequence that position. 10 nucleotides in length are identical to the corresponding positions in a second 10-nucleotide sequence, then the two sequences have 70% sequence identity. The length of comparison sequences will at least 50 nucleotides, preferably at least 60 generally be nucleotides, more preferably at least 75 nucleotides, and most preferably 100 nucleotides. Sequence identity is typically measured using sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705).

10

15

20

25

The present invention comprises a vector comprising a DNA sequence which encodes a human TADG-15 protein and said vector is capable of replication in a host which comprises, in operable linkage: a) an origin of replication; b) a promoter; and c) a DNA sequence coding for said protein. Preferably, the vector of the present invention contains a portion of the DNA sequence shown in SEQ ID No:1. A "vector" may be defined as a replicable nucleic acid construct, e.g., a plasmid or viral nucleic acid. Vectors may be used to amplify and/or express nucleic acid encoding TADG-15 protein. An expression vector is a replicable construct in which a nucleic acid sequence encoding a polypeptide is operably linked to suitable

control sequences capable of effecting expression of the polypeptide in a cell. The need for such control sequences will vary depending upon the cell selected and the transformation method chosen. Generally, control sequences include a transcriptional promoter and/or enhancer, suitable mRNA ribosomal binding sites, and which control the termination of transcription sequences translation. Methods which are well known to those skilled in the art can be used to construct expression vectors containing appropriate transcriptional and translational control signals. See for example, the techniques described in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual (2nd Ed.), Cold Spring Harbor Press, N.Y. A gene and its transcription control sequences are defined as being "operably linked" if the transcription control sequences effectively control the transcription of the gene. Vectors of the invention include, but are not limited to, plasmid vectors and viral vectors. Preferred viral vectors of the invention are those derived from retroviruses, adenovirus, adeno-associated virus, SV40 virus, or herpes viruses.

10

15

20

25

By a "substantially pure protein" is meant a protein which has been separated from at least some of those components which naturally accompany it. Typically, the protein is substantially pure when it is at least 60%, by weight, free from the proteins and other naturally-occurring organic molecules with which it is associated in vivo. Preferably, the purity of the naturally preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight. A substantially pure TADG-15 protein may be obtained, for example, by extraction from a natural recombinant by expression of nucleic acid a source;

encoding an TADG-15 polypeptide; or by chemically synthesizing the protein. Purity can be measured by any appropriate method, e.g., column chromatography such as immunoaffinity chromatography using an antibody specific for TADG-15, polyacrylamide gel electrophoresis, or HPLC analysis. A protein is substantially free of naturally associated components when it is separated from at least some of those contaminants which accompany it in its natural state. Thus, a protein which is chemically synthesized or produced in a cellular system different from the cell from which it naturally originates will be, by definition, substantially free from its naturally associated components. Accordingly, substantially pure proteins include eukaryotic proteins synthesized in *E. coli*, other prokaryotes, or any other organism in which they do not naturally occur.

In addition to substantially full-length proteins, invention also includes fragments (e.g., antigenic fragments) of the 15 TADG-15 protein (SEQ ID No:2). As used herein, "fragment," as applied to a polypeptide, will ordinarily be at least 10 residues, more typically at least 20 residues, and preferably at least 30 (e.g., 50) residues in length, but less than the entire, intact sequence. Fragments of the TADG-15 protein can be generated by methods 20 known to those skilled in the art, e.g., by enzymatic digestion of naturally occurring or recombinant TADG-15 protein, by recombinant DNA techniques using an expression vector that encodes a defined fragment of TADG-15, or by chemical synthesis. The ability of a candidate fragment to exhibit a characteristic of TADG-15 (e.g., 25 binding to an antibody specific for TADG-15) can be assessed by methods described herein. Purified TADG-15 or antigenic fragments generate new antibodies or to of TADG-15 can be used to

test existing antibodies (e.g., as positive controls in a diagnostic assay) by employing standard protocols known to those skilled in the art. Included in this invention are polyclonal antisera generated by using TADG-15 or a fragment of TADG-15 as the immunogen in, e.g., rabbits. Standard protocols for monoclonal and polyclonal antibody production known to those skilled in this art are employed. The monoclonal antibodies generated by this procedure can be screened for the ability to identify recombinant TADG-15 cDNA clones, and to distinguish them from known cDNA clones.

Further included in this invention are TADG-15 proteins which are encoded at least in part by portions of SEQ ID NO:2, e.g., products of alternative mRNA splicing or alternative protein processing events, or in which a section of TADG-15 sequence has been deleted. The fragment, or the intact TADG-15 polypeptide, may be covalently linked to another polypeptide, e.g. which acts as a label, a ligand or a means to increase antigenicity.

10

1,5

20

The invention also includes a polyclonal or monoclonal antibody which specifically binds to TADG-15. The invention encompasses not only an intact monoclonal antibody, but also an immunologically-active antibody fragment, e.g., a Fab or (Fab)₂ fragment; an engineered single chain Fv molecule; or a chimeric molecule, e.g., an antibody which contains the binding specificity of one antibody, e.g., of murine origin, and the remaining portions of another antibody, e.g., of human origin.

In one embodiment, the antibody, or a fragment thereof, may be linked to a toxin or to a detectable label, e.g. a radioactive label, non-radioactive isotopic label, fluorescent label, chemiluminescent label, paramagnetic label, enzyme

label, or colorimetric label. Examples of suitable toxins include diphtheria toxin, *Pseudomonas* exotoxin A, ricin, and cholera toxin. Examples of suitable enzyme labels include malate hydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, alcohol dehydrogenase, alpha-glycerol phosphate dehydrogenase, triose phosphate isomerase, peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase, acetylcholinesterase, etc. Examples of suitable radioisotopic labels include ³H, ¹²⁵I, ¹³II, ³²P, ³⁵S, ¹⁴C, etc.

5

10

15

20

25

Paramagnetic isotopes for purposes of in vivo diagnosis can also be used according to the methods of this invention. There are numerous examples of elements that are useful in magnetic resonance imaging. For discussions on in vivo nuclear magnetic resonance imaging, see, for example, Schaefer et al., (1989) JACC 14, 472-480; Shreve et al., (1986) Magn. Reson. Med. 3, 336-340; Wolf, G. L., (1984) Physiol. Chem. Phys. Med. NMR 16, 93-95; Wesbey et al., (1984) Physiol. Chem. Phys. Med. NMR 16, 145-155; Runge et al., (1984) Invest. Radiol. 19, 408-415. Examples of suitable fluorescent labels include a fluorescein label, an isothiocyalate label, a rhodamine label, a phycocrythrin label, a phycocyanin label, a n allophycocyanin label, an ophthaldehyde label, a fluorescamine label, Examples of chemiluminescent labels include a luminal label, an isoluminal label, an aromatic acridinium ester label, an imidazole label, an acridinium salt label, an oxalate ester label, a luciferin label, a luciferase label, an aequorin label, etc.

Those of ordinary skill in the art will know of other suitable labels which may be employed in accordance with

the present invention. The binding of these labels to antibodies or fragments thereof can be accomplished using standard techniques commonly known to those of ordinary skill in the art. Typical techniques are described by Kennedy et al., (1976) Clin. Chim. Acta 70, 1-31; and Schurs et al., (1977) Clin. Chim. Acta 81, 1-40. Coupling techniques mentioned in the latter are the glutaraldehyde method, the periodate method, the dimaleimide method, the m-maleimidobenzyl-N-hydroxy-succinimide ester method. All of these methods are incorporated by reference herein.

Also within the invention is a method of detecting TADG-15 protein in a biological sample, which includes the steps of contacting the sample with the labeled antibody, e.g., radioactively tagged antibody specific for TADG-15, and determining whether the antibody binds to a component of the sample.

10

15

20

25

As described herein, the invention provides a number of diagnostic advantages and uses. For example, the TADG-15 protein is useful in diagnosing cancer in different tissues since this protein is highly overexpressed in tumor cells. Antibodies (or antigen-binding fragments thereof) which bind to an epitope specific for TADG-15, are useful in a method of detecting TADG-15 protein in a biological sample for diagnosis of cancerous or neoplastic transformation. This method includes the steps of obtaining a biological sample (e.g., cells, blood, plasma, tissue, etc.) from a patient suspected of having cancer, contacting the sample with a labeled antibody (e.g., radioactively tagged antibody) specific for TADG-15, and detecting the TADG-15 protein using standard immunoassay techniques such as an ELISA. Antibody binding to the biological sample indicates that the sample contains a component which specifically binds to an epitope

within TADG-15.

10

15

20

25

Likewise, a standard Northern blot assay can be used to ascertain the relative amounts of TADG-15 mRNA in a cell or tissue obtained from a patient suspected of having cancer, in accordance with conventional Northern hybridization techniques known to those of ordinary skill in the art. This Northern assay uses a hybridization probe, e.g. radiolabelled TADG-15 cDNA, either containing the full-length, single stranded DNA having a sequence complementary to SEQ ID NO:1 (Figure 9), or a fragment of that DNA sequence at least 20 (preferably at least 30, more preferably at least 50, and most preferably at least 100 consecutive nucleotides in length). The DNA hybridization probe can be labeled by any of the many different methods known to those skilled in this art.

Antibodies to the TADG-15 protein can be used in an immunoassay to detect increased levels of TADG-15 protein expression in tissues suspected of neoplastic transformation. These same uses can be achieved with Northern blot assays and analyses.

The present invention is directed to DNA encoding a TADG-15 protein selected from the group consisting of: (a) isolated DNA which encodes a TADG-15 protein; (b) isolated DNA which hybridizes to isolated DNA of (a) above and which encodes a TADG-15 protein; and (c) isolated DNA differing from the isolated DNAs of (a) and (b) above in codon sequence due to the degeneracy of the genetic code, and which encodes a TADG-15 protein. Preferably, the DNA has the sequence shown in SEQ ID No:1. More preferably, the DNA encodes a TADG-15 protein having the amino acid sequence shown in SEQ ID No:2.

The present invention is also directed to a

vector capable of expressing the DNA of the present invention adapted for expression in a recombinant cell and regulatory elements necessary for expression of the DNA in the cell. Preferably, the vector contains DNA encoding a TADG-15 protein having the amino acid sequence shown in SEQ ID No:2.

5

10

20

The present invention is also directed to a host cell transfected with the vector described herein, said vector expressing a TADG-15 protein. Representative host cells include consisting of bacterial cells, mammalian cells and insect cells.

The present invention is also directed to a isolated and purified TADG-15 protein coded for by DNA selected from the group consisting of: (a) isolated DNA which encodes a TADG-15 protein; (b) isolated DNA which hybridizes to isolated DNA of (a) above and which encodes a TADG-15 protein; and (c) isolated DNA differing from the isolated DNAs of (a) and (b) above in codon sequence due to the degeneracy of the genetic code, and which encodes a TADG-15 protein. Preferably, the isolated and purified TADG-15 protein of claim 9 having the amino acid sequence shown in SEQ ID No:2.

The present invention is also directed to a method of detecting expression of the protein of claim 1, comprising the steps of: (a) contacting mRNA obtained from the cell with the labeled hybridization probe; and (b) detecting hybridization of the probe with the mRNA.

The following examples are given for the purpose of illustrating various embodiments of the invention and are not meant to limit the present invention in any fashion.

EXAMPLE 1

Tissue collection and storage

10

25

Upon patient hysterectomy, bilateral salpingo-oophorectomy, or surgical removal of neoplastic tissue, the specimen is retrieved and placed it on ice. The specimen was then taken to the resident pathologist for isolation and identification of specific tissue samples. Finally, the sample was frozen in liquid nitrogen, logged into the laboratory record and stored at -80°C. Additional specimens were frequently obtained from the Cooperative Human Tissue Network (CHTN). These samples were prepared by the CHTN and shipped to us on dry ice. Upon arrival, these specimens were logged into the laboratory record and stored at -80°C.

EXAMPLE 2

20 mRNA isolation and cDNA synthesis

Forty-one ovarian tumors (10 low malignant potential tumors and 31 carcinomas) and 10 normal ovaries were obtained from surgical specimens and frozen in liquid nitrogen. The human ovarian carcinoma cell lines SW 626 and Caov 3, the human breast carcinoma cell lines MDA-MB-231 and MDA-MB-435S, and the human uterine cervical carcinoma cell line Hela were purchased from the American Type Culture Collection (Rockville, MD). Cells were cultured to subconfluency in Dulbecco's modified Eagle's medium,

suspended with 10% (v/v) fetal bovine serum and antibiotics.

Messenger RNA (mRNA) isolation was performed according to the manufacturer's instructions using the Mini RiboSepTM Ultra mRNA isolation kit purchased from Becton Dickinson (cat. # 30034). This was an oligo(dt) chromatography based system of mRNA isolation. The amount of mRNA recovered was quantitated by UV spectrophotometry.

First strand complementary DNA (cDNA) was synthesized using 5.0 mg of mRNA and either random hexamer or oligo(dT) primers according to the manufacturer's protocol utilizing a first strand synthesis kit obtained from Clontech (cat.# K1402-1). The purity of the cDNA was evaluated by PCR using primers specific for the p53 gene. These primers span an intron such that pure cDNA can be distinguished from cDNA that is contaminated with genomic DNA.

15

10

EXAMPLE 3

PCR reactions

The mRNA overexpression of TADG-15 was determined using a quantitative PCR. Oligonucleotide primers were used for: TADG-15, forward 5'-ATGACAGAGGATTCAGGTAC-3' (SEQ ID NO: 10) and reverse 5'-GAAGGTGAAGTCATTGAAGA-3' (SEQ ID NO: 11); and β-tubulin, forward 5'-TGCATTGACAACGAGGC-3' (SEQ ID NO: 12) and reverse 5'-CTGTCTTGACATTGTTG-3' (SEQ ID NO: 13). β-tubulin was utilized as an internal control. Reactions were carried out as follows: first strand cDNA generated from 50 ng of mRNA will be used as template in the presence of 1.0 mM MgCl₂, 0.2 mM dNTPs, 0.025 U Taq polymerase/ml of reaction, and 1 x buffer supplied with

enzyme. In addition, primers must be added to the PCR reaction. Degenerate primers which may amplify a variety of cDNAs are used at a final concentration of 2.0 mM each, whereas primers which amplify specific cDNAs are added to a final concentration of 0.2 mM each.

After initial denaturation at 95°C for 3 minutes, thirty cycles of PCR are carried out in a Perkin Elmer Gene Amp 2400 thermal cycler. Each cycle consists of 30 seconds of denaturation at 95°C, 30 seconds of primer annealing at the appropriate annealing temperature, and 30 seconds of extension at 72°C. The final cycle will be extended at 72°C for 7 minutes. To ensure that the reaction succeeded, a fraction of the mixture will be electrophoresed through a 2% agarose/TAE gel stained with ethidium bromide (final concentration 1 mg/ml). The annealing temperature varies according to the primers that are used in the PCR reaction. For the reactions involving degenerate primers, an annealing temperature of 48°C were used. The appropriate annealing temperature for the TADG-15 and β-tubulin specific primers is 62°C.

20

25

10

EXAMPLE 4

T-vector ligation and transformations

The purified PCR products are ligated into the Promega T-vector plasmid and the ligation products are used to transform JM109 competent cells according to the manufacturer's instructions (Promega cat. #A3610). Positive colonies were cultured for amplification, the plasmid DNA isolated by means of the WizardTM Minipreps DNA purification system (Promega cat #A7500), and the plasmids were

digested with ApaI and SacI restriction enzymes to determine the size of the insert. Plasmids with inserts of the size(s) visualized by the previously described PCR product gel electrophoresis were sequenced.

5

10

15

EXAMPLE 5

DNA sequencing

Utilizing a plasmid specific primer near the cloning site, sequencing reactions were carried out using PRISMTM Ready Reaction Dye DeoxyTM terminators (Applied Biosystems cat# 401384) according to the manufacturer's instructions. Residual dye terminators were removed from the completed sequencing reaction using a CentrisepTM spin column (Princeton Separation cat.# CS-901). An Applied Biosystems Model 373A DNA Sequencing System was available and was used for sequence analysis. Based upon the determined sequence, primers that specifically amplify the gene of interest were designed and synthesized.

EXAMPLE 6

20

2.5

Northern blot analysis

10 μg mRNAs were size separated by electrophoresis through a 1% formaldehyde-agarose gel in 0.02 M MOPS, 0.05 M sodium acetate (pH 7.0), and 0.001 M EDTA. The mRNAs were then blotted to Hybond-N (Amersham) by capillary action in 20 x SSPE. The RNAs are fixed to the membrane by baking for 2 hours at 80°C. Additional multiple tissue northern (MTN) blots were purchased from CLONTECH Laboratories, Inc. These blots include the Human

MTN blot (cat.#7760-1), the Human MTN II blot (cat.#7759-1), the Human Fetal MTN II blot (cat.#7756-1), and the Human Brain MTN III blot (cat.#7750-1). The appropriate probes were radiolabelled utilizing the Prime-a-Gene Labeling System available from Promega (cat#U1100). The blots were probed and stripped according to the ExpressHyb Hybridization Solution protocol available from CLONTECH (cat.#8015-1 or 8015-2).

EXAMPLE 7

10 Quantitative PCR

15

20

Quantitative-PCR was performed in a reaction mixture consisting of cDNA derived from 50 ng of mRNA, 5 pmol of sense and antisense primers for TADG-15 and the internal control β-tubulin, 0.2 mmol of dNTPs, 0.5 mCi of [α-32P]dCTP, and 0.625 U of Taq polymerase in 1 x buffer in a final volume of 25 ml. This mixture was subjected to 1 minute of denaturation at 95°C followed by 30 cycles of denaturation for 30 seconds at 95°C, 30 seconds of annealing at 62°C, and 1 minute of extension at 72°C with an additional 7 minutes of extension on the last cycle. The product was electrophoresed through a 2% agarose gel for separation, the gel was dried under vacuum and autoradiographed. The relative radioactivity of each band was determined by PhosphoImager from Molecular Dynamics.

25

The present invention describes the use of primers directed to conserved areas of the serine protease family to

EXAMPLE 8

identify members of that family which are overexpressed in carcinoma. Several genes were identified and cloned in other tissues, but not previously associated with ovarian carcinoma. The present invention describes a protease identified in ovarian carcinoma. This gene was identified using primers to the conserved area surrounding the catalytic domain of the conserved amino acid histidine and the downstream conserved amino acid serine which lies approximately 150 amino acids towards the carboxyl end of the protease.

The gene encoding the novel extracellular serine protease of the present invention was identified from a group of proteases overexpressed in carcinoma by subcloning and sequencing the appropriate PCR products. An example of such a PCR reaction is given in Figure 1. Subcloning and sequencing of individual bands from such an amplification provided a basis for identifying the protease of the present invention.

EXAMPLE 9

The sequence determined for the catalytic domain of TADG-15 is presented in Figure 2 and is consistent with other serine proteases and specifically contains conserved amino acids appropriate for the catalytic domain of the trypsin-like serine protease family. Specific primers (20mers) derived from this sequence were used.

A series of normal and tumor cDNAs were examined to determine the expression of the TADG-15 gene in ovarian carcinoma. In a series of normal derived cDNA compared to carcinoma derived cDNA using β-tubulin as an internal control for PCR amplification, TADG-15 was significantly overexpressed in all of the

25

carcinomas examined and either was not detected or was detected at a very low level in normal epithelial tissue (Figure 3). This evaluation was extended to a standard panel of about 40 tumors. Using these specific primers, the expression of this gene was also examined in tumor cell lines derived from both ovarian and breast carcinoma tissues as shown in Figure 5 and in other tumor tissues as shown in Figure 6. The expression of TADG-15 was also observed in carcinomas of the breast, colon, prostate and lung.

Using the specific sequence for TADG-15 covering the full domain of the catalytic site as a probe for Northern blot analysis, three Northern blots were examined: one derived from ovarian tissues, both normal and carcinoma; one from fetal tissues; and one from adult normal tissues. As shown in Figure 7, TADG-15 transcripts were noted in all ovarian carcinomas, but were not present in detectable levels in any of the following tissues: a) normal ovary, b) fetal liver and brain, c) adult spleen, thymus, testes, overy and peripheral blood lymphocytes, d) skeletal muscle, liver, brain or heart. The transcript size was found to be approximately 3.2 kb. The hybridization for the fetal and adult blots was appropriate and done with the same probe as with the ovarian tissue. Subsequent to this examination, it was confirmed that these blots contained other detectable mRNA transcripts

10

. 15

20

25

Initially using the catalytic domain of the protease to probe Hela cDNA and ovarian tumor cDNA libraries, one clone was obtained covering the entire 3' end of the TADG-15 gene from the ovarian tumor library. On further screening using the 5' end of the newly detected clones, two more clones were identified covering the 5' end of the TADG-15 gene from the Hela library (Figure 8). The

complete nucleotide sequence (SEQ ID No:1) is provided in Figure 9 along with translation of the open reading frame (SEQ ID No:2).

In the nucleotide sequence, there is a Kozak sequence typical of sequences upstream from the initiation site of translation. There is also a poly-adenylation signal sequence and a polyadenylated tail. The open reading frame consists of a 855 amino acid sequence (SEQ ID No:2) which includes an amino terminal cytoplasmic tail from amino acids 1-50, an approximately 22 amino acid by an extracellular sequence transmembrane domain followed identified from complement CUB repeats preceding two subcomponents Clr and Cls. These two repeats are followed by four repeat domains of a class A motif of the LDL receptor and these four repeats are followed by the protease enzyme of the trypsin family constituting the carboxyl end of the TADG-15 protein (Figure 11). Also a clear delineation of the catalytic domain conserved histidine, aspartic acid, serine series along with a series of amino acids conserved in the serine protease family is indicated (Figure 10).

10

15

20

25

A search of GeneBank for similar previously identified sequences yielded one such sequence with relatively high homology to a portion of the TADG-15 gene. The similarity between the portion of TADG-15 from nucleotide #182 to 3139 and SNC-19 (SEQ ID No: 9; GeneBank accession #U20428) is approximately 97% (Figure 12). There are however significant differences between SNC-19 and TADG-15 viz. TADG-15 has an open reading frame of 855 amino acids whereas the longest ORF of SNC-19 is only 173 amino acids. SNC-19 does not include a proper start site for the initiation of translation nor does it include the amino terminal portion of the protein encoded by TADG-15. Moreover, SNC-19 does not include an ORF for a

functional serine protease because the His, Asp and Ser residues necessary for function are encoded in different reading frames.

TADG-15 is a highly overexpressed gene in tumors. It is expressed in a limited number of normal tissues, primarily tissues that are involved in either uptake or secretion of molecules e.g. colon and pancreas. TADG-15 is further novel in its component structure of domains in that it has a protease catalytic domain which could be released and used as a diagnostic and which has the potential for a target for therapeutic intervention. TADG-15 also has ligand binding domains which are commonly associated with molecules that internalize or take-up ligands from the external surface of the cell as does the LDL receptor for the LDL cholesterol complex. potential that these domains may be involved in uptake of specific ligands and they may offer the potential for making delivery of toxic molecules or genes to tumor cells which express this molecule on their surface. It has features that are similar to the hepsin serine protease molecule in that it also has an amino-terminal transmembrane domain with the proteolytic catalytic domain extended into the extracellular matrix. The difference here is that TADG-15 includes these ligand binding repeat domains which the hepsin gene does not have. In addition to the use of this gene as a diagnostic or therapeutic target in ovarian carcinoma and other carcinomas including breast, prostate, lung and colon, its ligand-binding domains may be valuable in the uptake of specific molecules into tumor cells.

10

1.5

20

25

Any patents or publications mentioned in this specification are indicative of the levels of those skilled in the art to which the invention pertains. These patents and publications are herein incorporated by reference to the same extent as

if each individual publication was specifically and individually indicated to be incorporated by reference.

One skilled in the art will readily appreciate that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. along with the methods, examples procedures, present The treatments, molecules, and specific compounds described herein are presently representative of preferred embodiments, are exemplary, and are not intended as limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art which are encompassed within the spirit of the invention as defined by the scope of the claims.

10

WHAT IS CLAIMED IS:

DNA encoding a TADG-15 protein selected from the
 group consisting of:

- (a) isolated DNA which encodes a TADG-15 protein;
- (b) isolated DNA which hybridizes to isolated DNA of (a) above and which encodes a TADG-15 protein; and
- (c) isolated DNA differing from the isolated DNAs of (a) and (b) above in codon sequence due to the degeneracy of the genetic code, and which encodes a TADG-15 protein.
- 2. The DNA of claim 1, wherein said DNA has the sequence shown in SEQ ID No:1.
 - 3. The DNA of claim 1, wherein said TADG-15 protein has the amino acid sequence shown in SEQ ID No:2.

20

4. A vector capable of expressing the DNA of claim 1 adapted for expression in a recombinant cell and regulatory elements necessary for expression of the DNA in the cell.

25

5. The vector of claim 4, wherein said DNA encodes a TADG-15 protein having the amino acid sequence shown in SEQ ID No:2.

WO 99/42120

6. A host cell transfected with the vector of claim 4, said vector expressing a TADG-15 protein.

5

7. The host cell of claim 6, wherein said cell is selected from group consisting of bacterial cells, mammalian cells, plant cells and insect cells.

10

8. The host cell of claim 7, wherein said bacterial cell is E. coli.

15

- 9. Isolated and purified TADG-15 protein coded for by DNA selected from the group consisting of:
 - (a) isolated DNA which encodes a TADG-15 protein;
- (b) isolated DNA which hybridizes to isolated DNA of (a) above and which encodes a TADG-15 protein; and
 - (c) isolated DNA differing from the isolated DNAs of (a) and (b) above in codon sequence due to the degeneracy of the genetic code, and which encodes a TADG-15 protein.
- 10. The isolated and purified TADG-15 protein of claim 9 having the amino acid sequence shown in SEQ ID No:2.

PCT/US99/03436

WO 99/42120

5

11. A method of detecting expression of the protein of claim 1, comprising the steps of:

(a) contacting mRNA obtained from the cell with the labeled hybridization probe; and

(b) detecting hybridization of the probe with the mRNA.

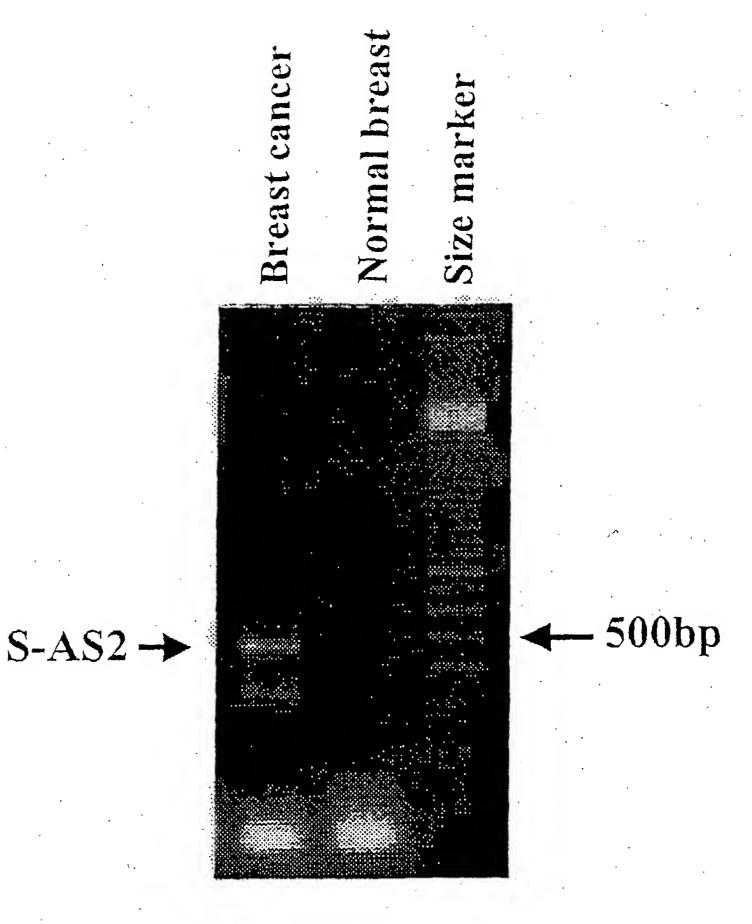


FIG. 1

VAQASPHGLC HDQSQRSAPC DTLGDR.F HNIEVLEG.P FDQGSDEE.P HDLSEHDGDF	NTQYYGQQ.1 HTQYGGTG.1 TTTSPDVTF1 NTASSGADY1 KTKYNANKT1 QLLDRGATA1 KHEALSPFY:	WRLCGIVSW(IFQAGVVSW(GTLQGLVSW(GQLQGVVSW(WYLYGIVSW(MYLTGIVSW(
LSRWRVFAGA PTQWTAFLGL MNEYTVHLGS KSRIQVRLGE RTSDVVVAGEAVLGE	GKICTVTGWG GKAIWVTGWG GTTCTVSGWG GTKCLISGWG GTLCATTGWG VRFSLVSGWG	CEDSISRTPRSSVEADGR CR CN CQKDGAHATHYRGT	Heps Tadg 15 Scce Try Chymb Fac 7 Tpa
PERNRV IDDRGFRYSD K Y DKIKNWRNLI QERFPPHHL.	AGQALVD ASHVFPA RCE PP APP AT ADDDFPA RTFSERTLAF LPPADLQLPD	COGDSGGPEV COGDSGGPL. CNGDSGGPLV CQGDSGGPLV CMGDSGGPLV CKGDSGGPLV CKGDSGGPLV	NO: 3) NO: 14) NO: 4) NO: 5) NO: 6) NO: 7)
DWVLTAAHCF NWLVSAAHCY RWVLTAAHC. QWVVSAGHC. DWVVTAAHC. IWVVSAAHCF CWILSAAHCF	EYIQPVCLPA SMVRPICLPD SMVKKVRLPS ARVSTISLPT QTVSAVCLPS DHVVPLCLPE QESSVVRTVC	GIDA GVDS KKNA GKDS GKDS GKDS	(SEQ. ID (SEQ. ID (SEQ. ID (SEQ. ID FP (SEQ. ID (SEQ. ID
HLCGGSLLSG HICGASLISP H.CGGVLVNE CGGSLINE HFCGGSLISE QLCGGTLINT FLCGGILISS	HLSS. PLPLT ELEK. PAEYS KLNS. QARLS KLSS. RAVIN KLAT. PARFS RLHQ. PVVLT QLKSDSSRCA	KMFCAGYPEG RMMCVGFLSG SMLCAGIPDS NMFCVGFLEG VMICAGAS YMFCAGYSDG	SEASGMVTQL ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
HALGQG LSGNQL NSGYHF QDKTGF LVNG. A	* EENSNDIALV DYDIALLHVNDLMLVLNNDIMLIVNNDITLL GTTNHDIALL DNDIALL	EYGN. QIKP LLPQ. QITP YKD. LLEN YPG. KITS WGR. RITD KVGDSPNITE LLNRT. VTD	EWIFQAIKTH DWIKENTGV~ KWINDTMKKH KWIKNTIAAN PWVQKILAAN EWLQKLMRSE DWIRDNMRP~
GEWPWQVSL. GEWPWQVSL. GSHPWQVAL. NSVPYQVSL. GSWPWQVSL. GECPWQVLL.	GYLPFRDPNS SHPFFNDFTF RHPGYSTQT. RHPQYDRKT. KNPKFSILT. PSTYVP	ISNDVCNGAD INQTTCEN ISPQDCTKV. LSQAKCEAS. LSNAECKKS. MTQDCLQQSR YPSSRCTSQH	GVYTRUSDER GVYTQVCKET GVYTKVYNYV GVYTRVSQYI GVYTRVSQYI GVYTRVSQYI
RIVGGRDTSL RVVGGTDADE KIIDGAPCAR KIVGGYNCEE RIVNGEDAVP RIVGGKVCPK RIVGGKVCPK	LGVQAVVYHG VQERRLKRII AQRIKASKSF EQFINAAKII IQVLKIAKVF QSRRVAQVII EQKFEVEKYI	GVLQEARVPI LILQKGEIRV SDLMCVDVKL DELQCLDAPV DKLQQAALPL ELMVLNVPRL ERLKEAHVRL	T. GCALAQKP D. GCAQRNKP D. GCAQKNKP SDTCS. TSSP Q. GCATVGHF
•			

ubulin — ADG15 — ADG15 —

normal ovary normal ovary m LMP a LMP s LMP a carcinoma s carcinom

normal ovary

FIG. 3

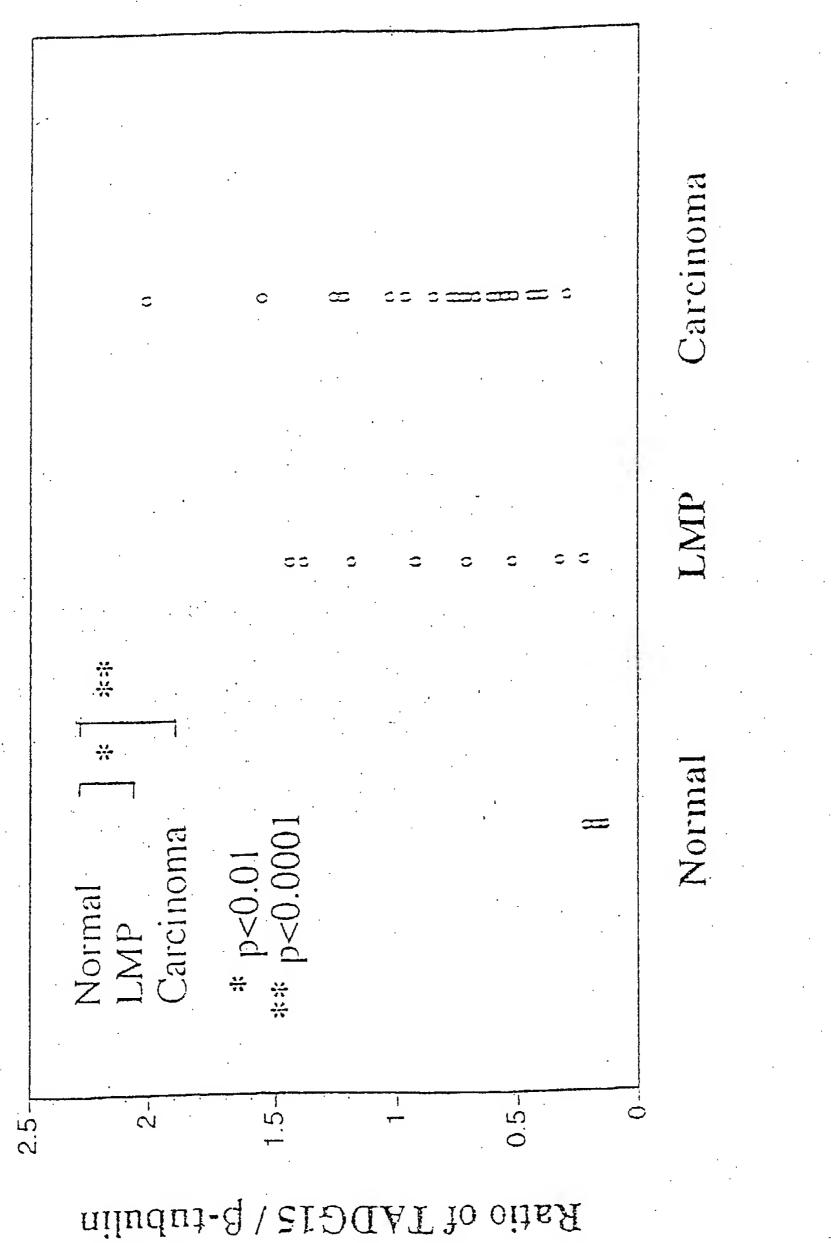


Figure 1

SUBSTITUTE SHEET (RULE 26)

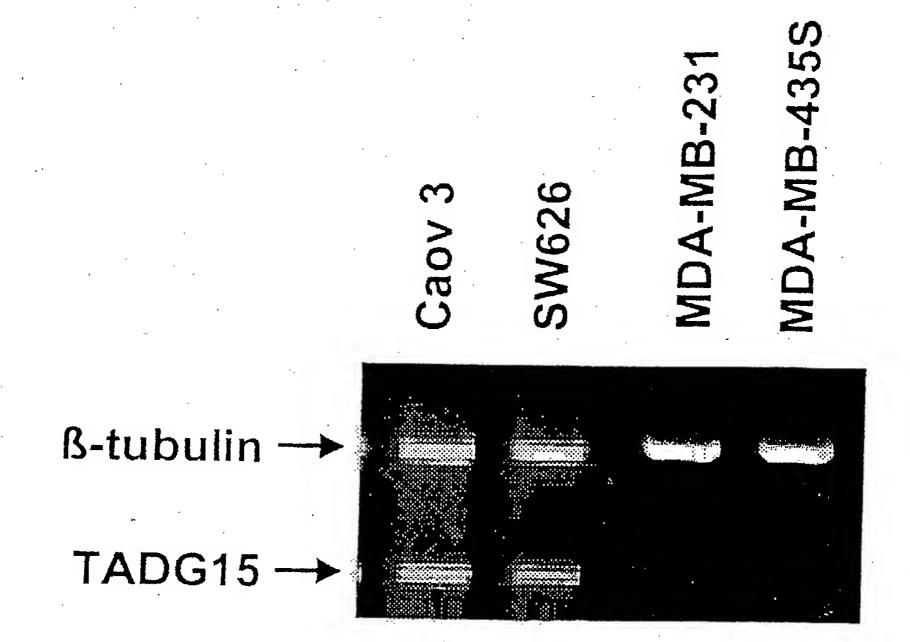


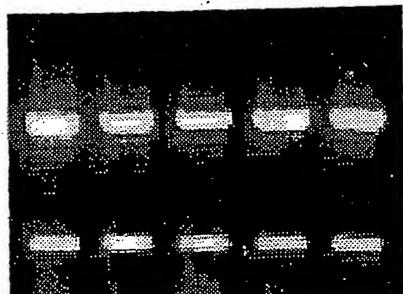
FIG. 5

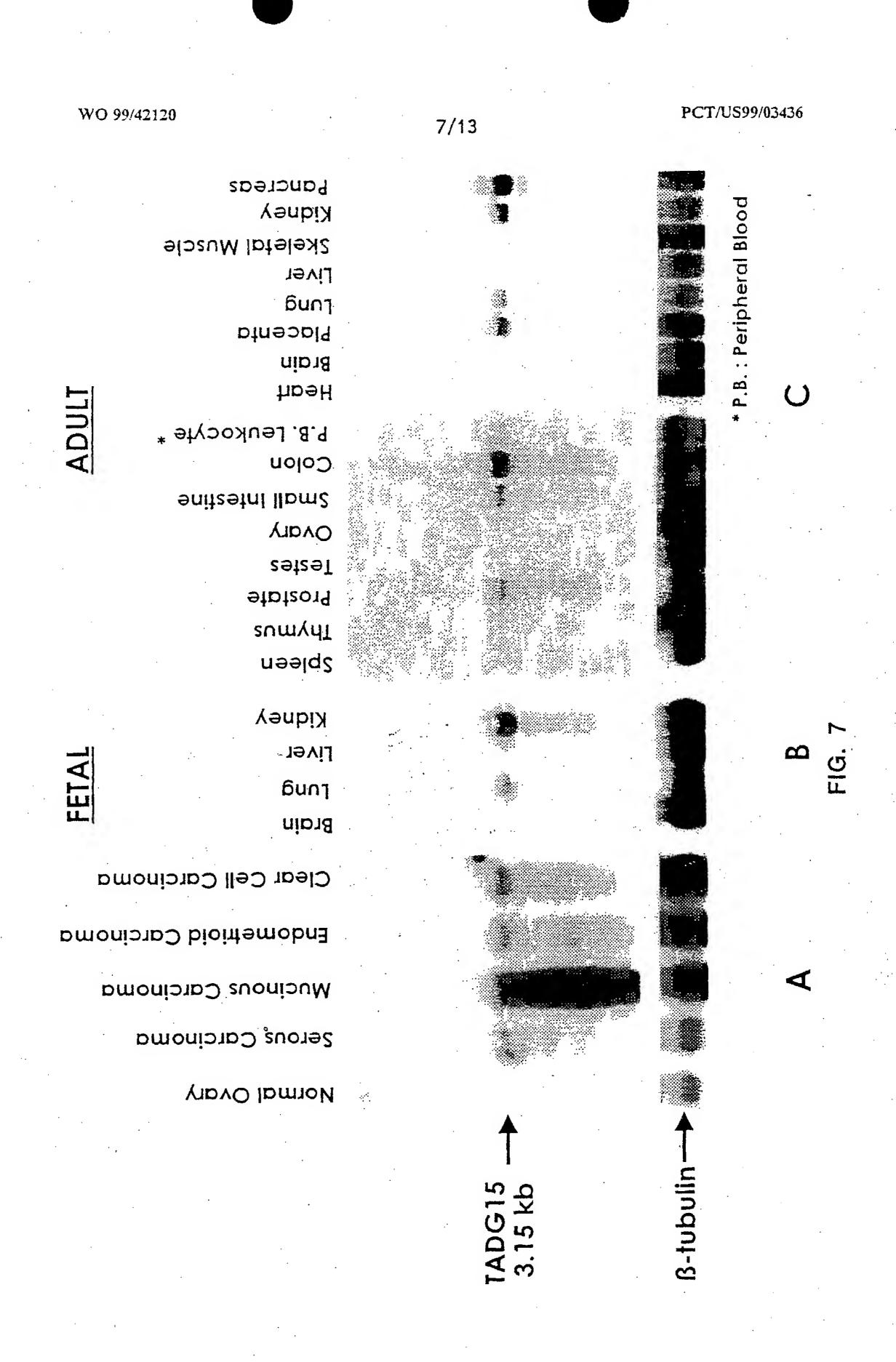
Ovarian cancer Breast cancer Colon cancer Prostate cancer

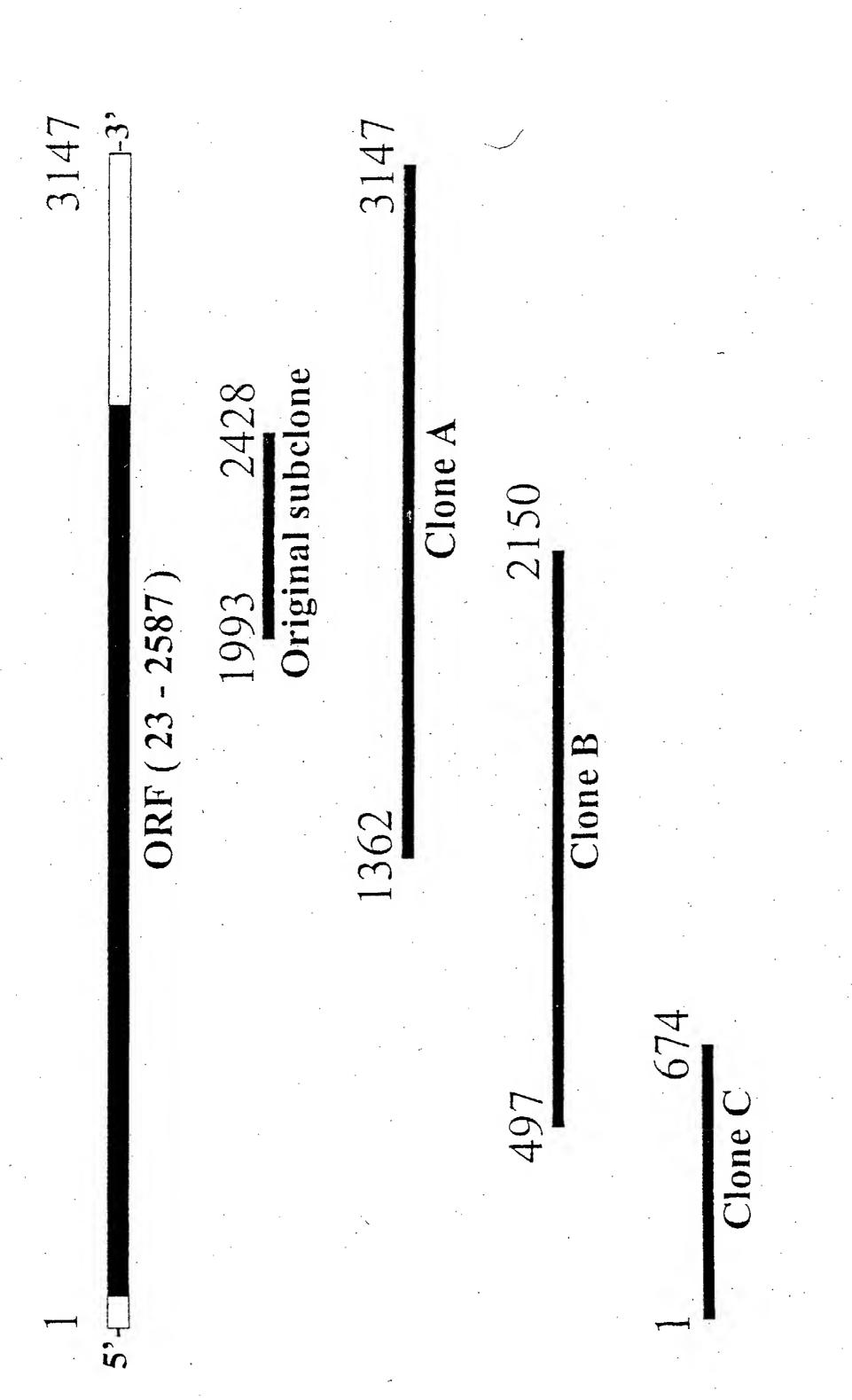
rostate cance Lung cancer

ß-tubulin →

TADG15 →







<u>اا</u> 2. ه

1 TONAGRAGOGGCCTCGGGGTACCATGGGGAGCGATCGGGCCCGCAAGGGCCGGAAGGACTTCGGCGCGGGACTCAAGTACAACTCCGGGCACGAGAAAGTGAATGGCTTGGA H G S D R A R K G G G F K D F G A G L K, Y N S R H E K V N G L E EGVEFLPVNNVKKVZKHGPGR<mark>WVVLAAVL</mark> 711 CTTCCTGGTGTGGCATTTGCAGTACCGGGACGTGCCAGAAGGTCTTCAATGGCTACATGAGGATCACAAATGAGAATTTTGTGGATGCCTACGAGAACTCCAACTCCACTGAGTT V W H L Q Y R D V R V Q K V F N G Y M R I T N E N F V D X Y E N S N S T E 361 TGTAAGCCTGGCCAGCAAGGTGAAGGACGCGCTGAAGCTGCTGTACAGCGGCAGTCCCATTCCTGGGCCCCTACCACAAGGAGTCGGCTGTGACGGCCTTCAGCGAGGGCAGCGTCATCCC V S L A S K V K D A L K L L Y S G V P T L G P Y H K E S A V T A F S E G S . T A Y Y W S E F S I P Q H L V E E A E R V M A E E R V V M L P P R A R S L K S F V V -TSVVAFFTDSKTVQRTQDNSCSFGLHARGVELMRFT.T?G? 731 CCCTGACAGCCCCTACCCCGCTCATGCCCGCTGCCAGTGGGCCCTGCGGGGGACGACGCCGACTCAGCTCAGCCTCACCTTCCGCAGCTTTGACCTTGCGTCCTGCGACGAGCGCGGCAG PDSPYPAHARCQWALRGDADSVLSLTFRSFDLAS D L V T V Y N T L S P H E P H A L V Q L C G T Y P P S Y N L 941 CATCACACTGATAACCAACACTGAGCGGCGGCATCCCGGCTTTGAGGCCACCTTCTTCCAGCTGCCTAGGATGAGCAGCTGTGGGGGCCGCTTACGTAAAGCCCAGGGCACATTCAACAG I T L I, T N T E R R H P G F E A T F F Q L P R M S S C G G R L R K A 1081 CCCCTACTACCCAGGCCACCCACCCAACATTGACTGCACATGGAACATTGAGGTGCCCAACAACCAGCATGTGAAGGTGAGCTTCAAATTCTTCTACCTGCTGGAGCCCGGCGTGCT PYTPGHTPPHIDCTHNIEVPHNQUVXVS 5 % K I T V R 5 H 5 D Q A G T C P K D Y V E I N G E K Y C G E R S Q F V V T S F SYTDIGELARY LSYDSSDEC PGOTTCRIGRCIRKELRCDS CTGCAAGCCCTTTTCTGGGTCTGCGACAGTSTGAACGA :::: CTGGGCCGACTGCACCGACACAGCGATGAGCTCAACTGCAGTTGCGACGCCGGCCAGTTCACGTGCAAGACAAGT WADCT BHS DELNCSCDAGHQFT CKNKFCKPLFW V CDS V N D 1111 CTGCGGAGACAACAGCGACGAGGAGGAGGAGTTGTCCGGCCCAGACCTTCAGGTGTTCCAATGGGAAGTGCCTCTCGAAAAGCCAGCAGTGCAATGGGAAGGACGACTGTGEGGACGA C G D N S D E Q G C S C P A Q T F R C S N G K C L S K S Q Q C N G K D D C G D G CAATGGGCTCTGCTTGAGCAAGGGCAACCCTGAGTGTGACGGGAAGGAGGACTG 151 GTCCGACGAGGCCTCCTGCCCCAAGGTGAACGTCGTCACTTGTACCAAACACACCCTACCGCTGCCT S D E A S C P K V N V V T C T K H T Y R C L N G L C L S K G N P E C D. G K E D C 1801 TAGOGRAGGETERGATGAGAGGRATGCGACTGTGGGCTGCGGTCATTCACGRARCAGGCTCGTGTTGTTGGGGGGCACGGATGCGGATGAGGCGAGGTGAGCCAGGTARGCCTGCR S D G S D E K D C D C G L R S F T R Q A R V V G G T D A D Z G E W P W Q V S L R 1921 TECTCTGGGCCAGGGCCACATCTGCGGTGCTTCCCTCATCTCTCCCAACTGGTCTCTGCCGCACACTGCTACATCGATGACAGGGTTCAGGTACTCAGACCCCACGTAGTGGTT ALGQGHICGASLISPN'W LVSAAH) CYID DRGFRYSDPTQW 2041 GSCCTTCCTGGGCTTGCACGACCAGAGCCAGCGCAGCGCCCCTGGGGTGCASGAGCGCAGGCTCAAGCGCATCATCTCCCACCCCTTCTTCAATGACTTCACCTTCGACTATEACATCGC AFLGLHDQSQRSAPGVQERRLKRIISEPFPN'DFTFDY©IA 2111 GETGE FGGAGETGGAGAAACCGGCAGAGTACAGETCCATGGTGCGGCCCATCTGCCTGCCGGACGCCTCCCATGTCTTCCCTGCCGGCAAGGCCATCTGGGTCACGGGCACACAC LLELEKPAEYSSMVRPICLPDASHVFPAGKAIWVTGMGHT 2291 CCAGTATGGAGGCACTGGCGCGCTGATCCTGCAAAAGGGTGAGATCCGCGTCATCAACCAGACCACCTGCGAGAACCTCCTGCGCAGATCACGCCGCGCATGATGTGEGTGGGCTT Q Y G G T G A L I L Q K G E I R V I N Q T T C E N L L P Q Q I T P R M M C Y G 2421 CCTCAGCGGCGGCGTTGGACTCCTGCCAGGGTGATTCCGGGGGACCCCTGTCCAGCGGAGGCGGATGGGGCGGATCTTCCAGGCCGGTGTGGAGACGGCTGCGGTCAGAG LSGGVDSCOGDGGPLSSVEADGRIFQAGVVSWGDGCAQR NKFGVYTRLPLFRDWIKENTGV 2842 CCCAGTGTGCACGCCTGCAGGCTGGAGACTGGACCGCTGACTGCACCAGCGCCCCCAGAACATACACTGTGAACTCAATCTCCAGGGCTCCAAAACCTCTCGCTTCCT 2312 COTTTGTGTATATCTGCCTCCCCTGTCTGTAAGGAGCAGCGGGAACGGAGCTTCGGAGCCTCCTCAGTGAAGGTGGTGGGGCCTGCGGATCTGGGGCCCTTGGTTCACGCTCT 2121 ITCITTTTAAAAAAAAAAAAAAA (SEQ IO NO: 1)

Li : Kozak's Concensus Sequence

O: Conserved amino acids of cutalytic triad H, D, S

Transmembrane domain

Figure 9.

1	MGSDRARKGG	GGPKDFGAGL	KYNSRHEKVN	GLEEGVEFLP	VNNVKKVEKH	1
51	GPGFWVVLAA	VLIGLLLVLL	GIGFLVWHLQ	YRDVRVQKVF	NGYMRITNEN	2
101	FVDAYENS	TEFVSLASKV	KDALKLLYSG	VPFLGPYHKE	SAVTAFSEGS	٠
151	VIAYYWSEFS	IPQHLVEEAE	RVMAEERVVM	LPPRARSLKS	FVVTSVVAFP	
201	TDSKTVQRTQ	DNSCSFGLHA	RGVELMRFTT	PGFPDSPYPA	HARČQWALRG	
251	DADSVLSLTF	RSFDLASČDE	RGSDLVTVYN	TLSPMEPHAL	VQLČGTYPPS	
301	YNLTFHSSQN	VLLITLITNT	ERRHPGFEAT	FFQLPRMSSC.	GGRLRKAQGT	3
351	FNSPYYPGHY	PPNIDČTWNI	EVPNNQHVKV	SFKFFYLLEP	GVPAGTČPKD	
401	YVEINGEKYČ	GERSQFVVTS	NSNKITVRFH	SDQSYTDTGF	LAEYLSYDSS	
451	DPCPGQFTCR	TGRCIRKELR	CDGWADCTDH	SDELNCSCDA	GHQFTCKNKF	
501	CKPLFWVCDS	VNDCGDNGDE	QGCSCPAQTF	RCSNGKCLSK	SQQCNGKDDC	4
551	GDGSDEASCP	KVNVVTCTKH	TYRCLNGLCL	SKGNPECDGK	EDCSDCSDEK	
601	DCDCGLRSFT	RQARVVGGTD	ADEGEWPWQV	SLHALGQGHI	CGASLISPNW	
651	LVSAAHCYID	DRGFRYSDPT	QWTAFLGLHD	QSQRSAPGVQ	ERRLKRIISH	
701	PFFNDFTFDY	DI ALLELEKP	AEYSSMVRPI	CLPDASHVFP	AGKAIWVTGW	5
751	GHTQYGGTGA	LILQKGEIRV	INQTTCENLL	PQQITPRMMC	VGFLSGGV:DS	Č
801	CQGCSGPLS	SVEADGRIFQ	AGVVSWGDGC	AQRNKPGVYT	RLPLFRDWIK	
851	ENTGV (SEQ.	.ID NO: 2)				•

: Conserved cysteine residue

: Possible N-linked glycosylation site

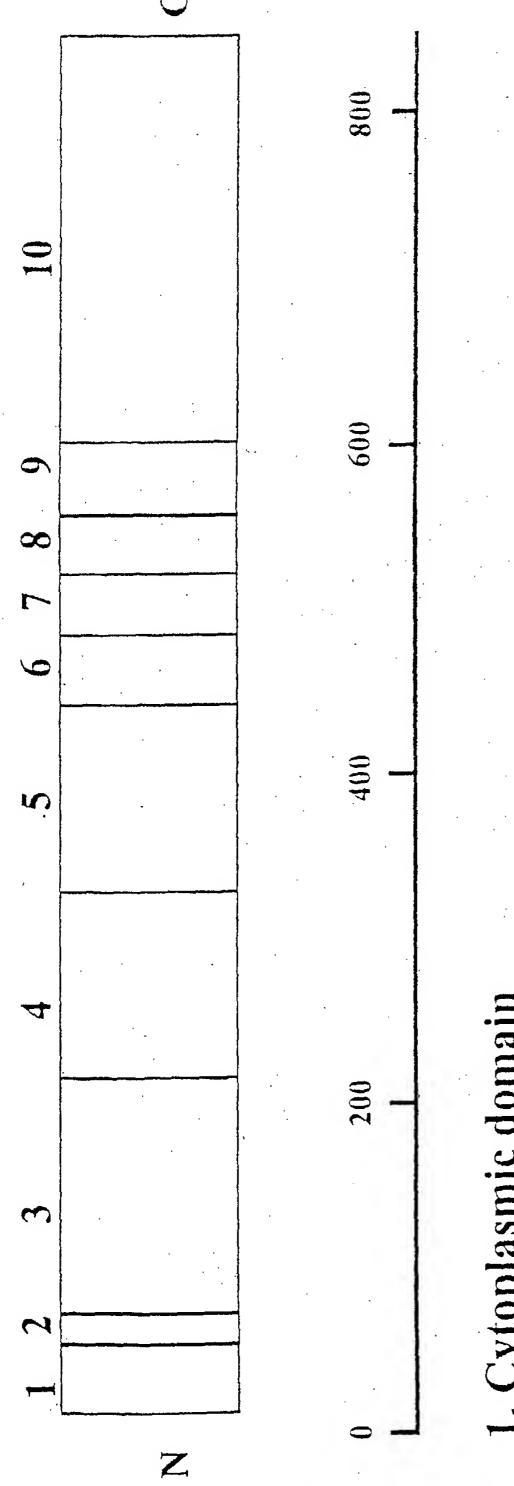
SDE: Conserved SDE motif

: Potential cleavage site

: Conserved amino acids of catalytic triad H, D, S

- 1. Cytoplasmic domain
- 2. Transmembrane domain
- 3. CUB repeat
- 4. Ligand-binding repeat (class A motif) of LDL receptor like domain
- 5. Serine protease

FIG. 10



Cytoplasmic domain
 Transmembrane domain

3. Extracellular domain

receptor like domain 6-9. Ligand-binding repeat (c 4-5. CUB repeat

10. Serine protease

FIG. 11

12/13

	LOCUS		2900 bp			57.3	17-MAR-19	9 7		-
	DEFINITIO ACCESSION		WKKW Sednen							
	HID KEYWORDS	diaaar.		•			v			
	SOURCE ORGENISM	Homo sapiens	irochondrial	eukaryct	45) Meta	ics: Chor	da ta :			
	Reference	Vertebrata: E	otheriz: Pri	.maces: Ca	15,777.171	: 2227472	14; X-444.			
	AUTHORS TITLE	Zheng, S., Cai SNC19 gene in	X., Geng, L.	, Cao,J., is	i∴≞∴ạ, L	. and thi	, 2 , 2 -			
	JOURNAL REFERENCE	Unpublished 2 (bases 1				·			•	
	AUTHORS TITLE	Theng, S. Direct Submis	sion	the Thoma			ianiana		:	
	JOURNAL	Submitted (30: Medical Univer	-JAN-1995) 5 rsicy, Hangz	hou. 3100	13B1	les Republ	iis of Chi	1.3		
	•									
	, 7	ADG15: TCAAGAG	GGCCTCGGGGT	ACCATGGEG	a 302a.TCG	3500000446	1110040001	300CCG-4GGA	CTTCGGCGCGCACT	.81
		:C19:								107
. 9:	2 CAAGTACAACTC	ICGGCACGAGAAAG:	<u>rgaat</u> ggcttgg	2729AADDA		CTGCIAGTG	and the second		AAGCATGCCCGGGG	
							-:::::::::::::::::::::::::::::::::::::	TTGCAGTACC	GGGACGTGCGTGTCC	281
187	CGCTGGGTGGTGG	TEGCAGCEGTECT	EATCGGCCICCI			11111111	111111111 1351GTG3CA	TTTGCAGTACC	GGCACGTGCGTGTCC	100
				**********			-1107004010	gasttistaag	CCTGGCCAGCAAGGT	
	2 AGEAGGICIIG 	HIGGUIACAIGAGE HILLIHIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII			: ::: 		HIHITIHI CAACTOCACT	HIIIIIIII Agritgiaag	CTGGCCAGCAAGGT	200
. W :				הררשההמלה	ومروم وحجام		اعتامية المسامة	DEDDORDITO	GGCAGCGTCATCGCC	481
201	CARGGACGCGCTC	HAGCTGCTGTACAG		TCCTGGGSC		i: (:::!!!!! PAGEAGTCG:	ctstgacsg	CTTCAGCGAG	HILLIHILI IIIIIIIIIIIIIIIIIIIIIIIIIIIII	300
; 8 ;	TACTACTGGTCTC	EAGTTCAGCATCCC	CAGCACCTGGT	GG#GG#GGC	constate		::::::::::::::::::::::::::::::::::::::	ASTCATGCTGC 1	CCCCGCGGGCGCGCT	581
301	TROTACTOGTCT(LAGTTCAGCATCCC(CAGCACCTGGT	المرابحيناتان.	ند شحرمتان نتال لي.	د ر سال المان				399
382	2 CCCTGAAGTCCTT	TGTGGTCACCTCAC	TGGTGGCTTTC	CCCACGGAC	TOCAAAACI	rstrcagas:	eaccageac Hilliiii	vacagetgeag	CTTTGGECTGCAEGC	153
1400) cocreaagrocr:	TETESTCACCICAC	TGGTGGC111C	المريسية المراساته المراسات		ه ندا طباه کید. پی در	2 (44			761
63.3	CCGCGGTGTGGAS	CTGATGCGCTTCAC	CACGCCCGGCT	TOCOTGRON	SICCOTAC:	::CG:TCAT	;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;		CGGGGGGACGCGAC	592
499) CCGCGGTGTGGA(CTEATGCGCTTCA(C2C2. CCGGC1				://	-ceetetacaa	TACCAGACA, CARC	876
793	CORRECTOR STATE OF ST	TCACCTTC	GCAGCTTTGAC	11 1	: i		111111111	II IIIIIIII AC GTGTACAA	CACCCTGAGCCCCAT	656
593		·					- ^~~~~~	COTTOTION	TCATCACACTGATAA	974
577	י פנייפרנרביני יוון פנייפרנרביני	CIGGIGCAGIIGIC	11111111111111111111111111111111111111	CTCCCTCCT	::::::!!! RCAACCTG	HILLIHI ACCTTCCACT	I IIIIIII TOOCTOOCAE	HIIIIIIII BACGTCCTGC	TCATCACACTGATAA	783
975								CCTTACGTAA	ASCCCAGGGGACATT	1074
764			TTGAGGCCACC'	TTCTTCCAG	TTGCCTAG:	eat gagerg	1999-990	CGCTTACGTAA	AGCCCAGGGGACATT	881
1075	CRACAGOCCCTAC	TACCCAGGCCACTA	.CCCACCCAACA	TTGRCTGCR		rtt <u>e</u> rggtg:		ACCATOTGAAG	GTGAGCTTCAAATTC	1174
892	CAACAGCCCCTAC	TACCCAGGCCACTA	CCCACCC-ACA	المراجعة المراجعين المراجعين		270010,	_041			981
1175	TICTACCTGCTG	AGCCCGGCGTGCCT	GCGGGCACCTG	CCCCAAGGA	CTACOTGO 	AGATCAATG:	33 <u>7464 A</u> ATA 	otgografia 	GGTCCCAGTTCGTCG 	1081
982	: TTETACCTGCTGC	lagecegecetecet	CCCCCCCCC	الدار تسميم الواحدة		مريبي والمريض والمعيني	٠٠٠٠٠ هـ عنااتيد ټول			1374
1275	TCACCAGCAACAC	CARCARGATCACAC 	TTCGCTTCCAC	TCAGATCAS	TUCTRCAC(544. TACCGG:			CTACGACTCCAGTGA 	1161
1082	TCACCAGCAACAC	CAACAAGATCACA(TTCGCTTCCAC	TCALA: Lac			tantaaccc tantaacccc	actecacegac	CACAGOGATGAGGTC	1474
1375	CCCATGCCCGGG	CASITCACGTGCCC 		GIAICCOUR GTATCCGC=	AGGAGGTA	111;11111. 130737347	::		HIHIHHHHHH CACAGCCATCAGCTC	1290
1152		•		~ * * * * * * * * * *			-557676767	<u> </u>	ACTGCGGAGACAACA	1574
1797	0001100010001000 			 GAGCALGTT	1::: TT3CARG.	HIIIII ctcttct:	illililli Gagtetgeba	HIIIIIIIII CAGTOTGAACG	AGTGCGGAGACAACA	1377

	THE STANDARD SET OF THE PROPERTY OF THE PROPERTY OF THE STANDARD SET OF THE STANDARD S	1673
1575	GCGACCAGCAGGGGGGGCAGTTGTCCGG.CCCAGACCTTCAGGTGTTCCAATGGGAAGTGCCTCTCGAAAAGCCAGCAGTGCAATGGGAAGGACGACTGTG	1477
1378	AARIA CII GAATTA AAAAAAA AAAAA AAAAAA AAAAAAAAAA	1773
1674	GGGACGGGTCCGACGAGGGCTCCTGCCCCAAGGTGAACGTCGTCACTTGTACCAAACACACCTACCGCTGCCTGAATGSGTTCTGCTTGAGCAAGGGCAA 	1677
1279	RESPONDE DE LA CONTRACTOR DE LA CONTRACTA DE LA CALACACACACACACACACACACACACACACACACA	1511
1771	COUTGRATGEGACGGGRAGGAGGACTGTAGCGACGGCTCAGATGRGPAGGACTGCGACTGTGGGCTGATTTTAGGAGACAGGCTCGTGTTGTTGG	1817
1570	CCCTGAGTGTGACGGGAAGGAGGACTGTAGCGACGGCTCAGATGAGAAGGACTGCGACTGTGGGCTGCGGCTGAGTGACGACACGAGACAGGCTCGTGTTGTTGGG LILILILILILILILILILILILILILILILILILILI	1677
:376	DDFTSALDDFTSTCTCCTCATCTCCCATCTCCCCACCCCCCCCCCCCCC	1973
18/5	GGCACGGATGCGGATGAGGGCGAGTGGCCCTGGCAGGTAAGCCTGCATGCTCTGGGCCAGGGCCACATCTGCGGTGCTTCCCCAACTGGC	1777
1678	THE SECOND CONTROL OF THE SECOND CONTROL OF	2073
1974	TOGTCTCTGCCGCACACTGCTACATCGATGACACGCATTCAGGTACTCAGGTACTCACGCACG	1875
1779	TGGTCTCTGCCGCACACTGCTACATCGATGACAGGGATTCAGGAATTCAGGAATTGAGGAATTGACATCGACTATGACATCGCGCTGCTGGAGGTG	2173
2074	CAGCGCCCCTGGGGTGCAGGAGGGCAGGCTCAAGCGCATCATCTCCCCACCCCTTCTTCAATGACTTCACCTTCGACTATGACATCGCGCTGCTGGAGCTG LI IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	1973
1876	ra recetedesestedada de la composição de	****
2174	GAGALACCGGCAGAGTACAGCTCCATGGTGCGGCCCATCTGCCTGC	2213
1974	CACESACCESCACACTACACECTECATUS I CUCCUCCAT CI OCCI TOCCI TOCC	2073
7774	GACACACCAGTATGGAGGCACTGGCGCGCTGATCCTGCAAAAGGGTGAGATCCGCGTCATCAACCAGACCACCTGCGAGAACCTCCTGCGCAGCAGCAGCACCAC	2373
	THE FIRST PROPERTY OF THE PROP	2173 2473
フスプュ	CACCCCCCATGATGTGTGGGCTACCTCAGCGGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCCATGATGTGTGTG	
2174		7577
2474	ATCTTCCAGGCCGGTGTGGTGAGCTGGGGAGACGGCTGCGCTCAGAGGAACAAGCCAGGCGTGTACACAAGGCTCCCTGTGTTTCGGGACTGGATCAAAG ATCTTCCAGGCCGGTGTGGTGAGCTGGGGAGACGGCGCGCTCAGAGGAACAAGGCCAGGCGTGTACACAAGGCTCCCTGTGTTCGGGAACGATCAAAG	2313
2274	TOTAL CONCENTRATES OF SECTEMBERS OF SECTION OF THE	2342
2574	ACAL CACTOGOGY AT AGGGGCCGGGGCCACCCAAATGTGTACACCTGCGGGGCCACCCATCGTCCACCCCAGTGTGCACGCCTGCAGGCTGGAGACT	2010
7777	AGAACACTGGGGTATAGGGGCCGGGGCCACCAAATGTGTACACGTGCGGGGCCACCCATCG.CCACGCCAGTGTGCACGCCTGCAGGCTGGAGACTCGC 111111111111111111111111111111111	2472
22/3	TOTAL	2770
2671	GGACCGCTGACTGCACCAGCGCCCCAGAACATACACTGTGAACTCAATCTCCAGGGCTCCAATCTGCCTACAATCTGCCTACAATCTGCCTACAATCTGCCTACAATCTGCCTACAATCTGCCTAGAAACCTCTGGCTTCCTCAGCCTCCAACCACGTGACCTGCACCACCAGCAACATCTGCAACCTCTCAGCCTCCAACCACCTGTGAACCTCTCAGCCTCCAACCACCTGTGAACCTCTCAGCCTCCAACCTGTGAACCTCTCAGCCTCCAACCTGTGAACCTCTCAGCCTCCAACCTGTGAACCTCTCAGCCTCCAACCTGTGAACCTCTCAGCCTCCAACCTGTGAACCTCTCAGCCTCCAACCTGTGAACCTCTCAGCCTCCAACCTGTGAACCTCTCAGCCTCCAACCTGTGAACCTCTCAGCCTCCAACCTGTGAACCTCTCAGCCTCCAACCTGTGAACCTCTCAGCCTCCAACCTGTGAACCTCTCAGCCTCCAACCTGTGAACCTGTGAACCTCTCCAGGCTCCAACCTGTGAACCTCTCCAGGCTCCAACCTGTGAACCTCTCCAGGCTCCAACCTGTGAACCTCTCCAGGCTCCAACCTGTGAACCTCTCAACCTGTGAACCTCTCCAGGCTCCAACCTGTGAACCTGTGAACCTCTCAACCTGTGAACCTCTCCAGGCTCCAACCTGTGAACCTGTGAACCTGTGAACCTCTCCAGGCTCCAACCTGTGAACCTGAACCTGTGAACCTGTGAACCTGTGAACCTGTGAACCTGTGAACCTGTGAACCTGTGAACCTGTGAACCTGTGAACCTGAACACACAC	2567
2473		2368
2771	AGTGGAGCTGGGA.GGTAGAAGGGGAGG.ACACTGGTGGTTCTACTGACCCAACTGGGGG.CAAGGTTTGAAG.CACAGGTCCGGCAGCCCAAG	2659
2568	ACTGCACCTGCCAGGGTAGAAGGGGAAGCACACACACACA	2955
2869	TGGGCCGAGGCGCGTTTGTGTATATCTGCCTCCCCTGTCTGT	2735
2659	TGGGCCGAGGCGCGTTTGTGTATATCTGCCTCCCCTGTCTGT	3065
3060	ANTERESE SE	,,,,,
	THE TENENT OF THE PROPERTY OF	2024
3060	THE SECTION CONTRACTOR STATES OF A LANGE OF	
2003	CCCAGGGTGGACTTCAGTGTGTATTTGTGTAAATGGGTAAACATTTATTT	9)

Figure 12 (cont.)

SEQUENCE LISTING

<110> O'Brien, Timo Tanimoto, Hir	77
	xtracellular Serine Protease l in Breast and Ovarian Carcinomas
<130> D6064PCT	•
<140> PCT/US99/0343	6
<141> 1999-02-18	
<150> US 09/027,337	
<151> 1998-02-20	
<160> 14	
<210> 1	
<211> 3147	
<212> DNA	
<213> Homo sapiens	
<220>	
<223> cDNA sequence	of TADG-15
<400> 1	
tcaagagcgg cctcggggta ccatggg	gag cgatcgggcc cgcaagggcg 50
gagggggccc gaaggacttc ggcgcgg	
gagaaagtga atggcttgga ggaaggc	gtg gagttcctgc cagtcaacaa 150
cgtcaagaag gtggaaaagc atggccc	ggg gegetgggtg gtgetggeag 200
ccgtgctgat cggcctcctc ttggtct	tgc tggggatcgg cttcctggtg 250
tągcatttgc agtaccggga cgtgcgt	gtc cagaaggtct tcaatggcta 300
catgaggatc acaaatgaga attttgt	gga tgcctacgag aactccaact 350
ccactgagtt tgtaagcctg gccagca	agg tgaaggacgc gctgaagctg 400
ctgtacagcg gagtcccatt cctgggc	ccc taccacaagg agtcggctgt 450
gacggccttc agcgagggca gcgtcat	cgc ctactactgg tctgagttca 500
gcatcccgca gcacctggtg gaggagg	ceg agegegteat ggeegaggag 550
cgcgtagtca tgctgccccc gcgggcg	gc tccctgaagt cctttgtggt 600
cacctcagtg gtggctttcc ccacgga	ctc caaaacagta cagaggaccc 650
aggacaacag ctgcagcttt ggcctgc	acg cccgcggtgt ggagctgatg 700
cgcttcacca cgcccggctt ccctgaca	age ecctaceeg etcatgeeeg 750
ctgccagtgg gccctgcggg gggacgc	ga ctcagtgctg agcctcacct 800
teegeagett tgacettgeg teetgega	acg agegeggeag egacetggtg 850
acggtgtaca acaccctgag ccccatgg	gag ccccacgccc tggtgcagtt 900
gtgtggcacc taccctccct cctacaac	ect gacettecae tecteceaga 950
acgtcctgct catcacactg ataaccaa	aca ctgagcggcg gcatcccggc 1000

tttgaggcca	ccttcttcca	gctgcctagg	atgagcagct	gtggaggccg	1050
			cccctactac		
			ttgaggtgcc		
catgtgaagg	tgagcttcaa	attcttctac	ctgctggagc	ccggcgtgcc	1200
			gatcaatggg		
gcggagagag	gtcccagttc	gtcgtcacca	gcaacagcaa	caagatcaca	1300
			gacaccggct		
atacctctcc	tacgactcca	gtgacccatg	cccggggcag	ttcacgtgcc	1400
gcacggggcg	gtgtatccgg	aaggagctgc	gctgtgatgg	ctgggccgac	1450
tgcaccgacc	acagcgatga	gctcaactgc	agttgcgacg	ccggccacca	1500
gttcacgtgc	aagaacaagt	tctgcaagcc	cctcttctgg	gtctgcgaca	1550
gtgtgaacga	ctgcggagac	aacagcgacg	agcaggggtg	cagttgtccg	1600
gcccagacct	tcaggtgttc	caatgggaag	tgcctctcga	aaagccagca	1650
gtgcaatggg	aaggacgact	gtggggacgg	gtccgacgag	gcctcctgcc	1700
ccaaggtgaa	cgtcgtcact	tgtaccaaac	acacctaccg	ctgcctcaat	1750
gggctctgct	tgagcaaggg	caaccctgag	tgtgacggga	aggaggactg	1800
tagcgacggc	tcagatgaga	aggactgcga	ctgtgggctg	cggtcattca	1850
cgagacaggc	tcgtgttgtt	gggggcacgg	atgcggatga	gggcgagtgg	1900
ccctggcagg	taagcctgca	tgctctgggc	cagggccaca	tctgcggtgc	1950
ttccctcatc	tctcccaact	ggctggtctc	tgccgcacac	tgctacatcg	2000
atgacagagg	attcaggtac	tcagacccca	cgcagtggac	ggccttcctg	2050
ggcttgcacg	accagagcca	gcgcagcgcc	cctggggtgc	aggagcgcag	2100
gctcaagcgc	atcatctccc	accccttctt	caatgacttc	accttcgact	2150
atgacatcgc	gctgctggag	ctggagaaac	cggcagagta	cagctccatg	2200
gtgcggccca	tctgcctgcc	ggacgcctcc	catgtcttcc	ctgccggcaa	2250
ggccatctgg	gtcacgggct	ggggacacac	ccagtatgga	ggcactggcg	2300
cgctgatcct	gcaaaagggt	gagateegeg	tcatcaacca	gaccacctgc	2350
gagaacctcc	tgccgcagca	gatcacgccg	cgcatgatgt	gcgtgggctt	2400
cctcagcggc	ggcgtggact	cctgccaggg	tgattccggg	ggacccctgt	2450
ccagcgtgga	ggcggatggg	cggatcttcc	aggccggtgt	ggtgagctgg	2500
ggagacggct	gcgctcagag	gaacaagcca	ggcgtgtaca	caaggeteee	2550
tctgtttcgg	gactggatca	aagagaacac	tggggtatag	gggccggggc	2600
cacccaaatg	tgtacacctg	cggggccacc	catcgtccac	cccagtgtgc	2650
acgcctgcag	gctggagact	ggaccgctga	ctgcaccagc	gcccccagaa	2/00
catacactgt	gaactcaatc	tccagggctc	caaatctgcc	tagaaaacct	2750
ctcgcttcct	cagcctccaa	agtggagctg	ggaggtagaa	ggggaggaca	2800
ctggtggttc	tactgaccca	actgggggca	aaggtttgaa	gacacagcct	2850
ccccgccag	ccccaagctg	ggccgaggcg	cgtttgtgta	tatctgcctc	∠ 900

	CCC	ctgt	ctgt	aag	gagca	agc q	ggga	acgga	ag c	ttcg	gagc	c tc	ctca	gtga	2950	}
	agg	gtgg	rāāā	gct	gccgg	gat d	tgg:	gctgi	tg gg	ggcc	cttg	g gc	cacg	ctct	3000)
	tga	aggaa	agcc	cag	gctc	gga g	gac	cctgg	ga aa	aacag	gacg	g gto	ctgag	gact	3050	•
	gaa	aattq	gttt	taco	cagct	icc d	cagg	gtgga	ac t	cagt	gtgt	gta	attt	gtgt	3100	-
	aaa	ıtggg	ytaa	aaca	aattt	at t	stati	tttt	aa aa	aaaa	aaaa	a aaa	aaaa	3.	3147	
		. •	21.05		3											
			210>	-	2									·		
			211> 212>		855 PRT											
	·		213>			5 C 31	nian	c	-	•					٠.	
			220>		LIOIN	J Sal	pien	ప								
			223>		Amir	70 P	-ia	cemia	ence	of s	בעומיו.	_15 4	271 <i>C</i> *()/	∂⊝∂ 1	by cI	ስ አ ፖማ
			100>		2	io ac	ناسب√ساست	sedra		<u> </u>	الملك	ا لىد		ueu .	oà cr	MA.
											-		-			
	Met	Gly	Ser	Asp	Arg	Ala	Arg	Lys	Gly		· Gly	Gly	Pro	Lys	Asp	
					5				٠	10					15	
,	Phe	Gly	Ala	Gly		Lys	Tyr	Asn	Ser			Glu	Lys	Val	Asn	
	 3		4.7		20		- 3		_	25			_		30	
	GTA	Leu	Glu	Glu			GLu	Phe	Leu			Asn	Asņ	. Val	•	•
	· ·	**- 7	~ 3	_	35			~ 1		40		T7 - 7	T	3. 7.	45 -	:
٠	rys	val	Glu	ГÀЗ		GTA	bro	GTĀ	Arg		Val	Var	Leu	Ala		
	₹ <i>7-</i> ~ 7	Τ	T1 -	01	.50	 T 0	T 0	TT- 7	т о	55	C 1	~1 ~	01	Dla a	60	,
	· var	ren	Ile	GTÅ	ьеи 65	теп	TIEU	Val	rea	70.		TTE	GTĀ	rne	75	,
	T a T	سدن	His	Ī (Al)		('' ''	222	A en	77= 7			Gln	Tare	175 1		
	۷ ئىنى⊥		1173	Tien	80		mir y	ಬಾಗಿ	A CT.T.	85	,v ct.r	G-77.1	my 2	val	90	
	Asn	Glv	Tyr	Met		Tle	ጥኮታ	Asn	Glu		Phe	Val	Asn	Ala		
		~_ ₁	· • · · ·	2300	95	4 • •	444			100		, az	-122		105	
	Glu	Asn	Ser	Asn		Thr	Glu	Phe	Val		Leu	Ala	Ser	Lvs		
î					110					115		•			120	
	Lys	Asp	Ala	Leu	Lys	Leu	Leu	Tyr	Ser	Gly	Val	Pro	Phe	Leu		
					125		-	_		130	•				135	
	Pro	Tyr	His	Lys	Glu	Ser	Ala	Val	Thr	Ala	Phe	Ser	Glu	Gly	Ser	
					140					145					150	
	Val	Ile	Ala	Tyr	Tyr	Trp	Ser	Glu	Phe	Ser	Ile	Pro	Gln	His	Leu	•
					155					160					165	
	Val	Glu	Glu	Ala	Glu	Arg	Val	Met	Ala	Glu	Glu	Arg	Val	Val	Met	
					170					175					180	
												•				

						•								
Leu	Pro	Pro	Arg			Ser	Leu	Lys		Phe	Val	Val	Thr	Ser
				185		_	_		190	rr_ 1		7. ***	mb~	•
Val	Val	Ala	Phe		Thr	Asp	Ser	rys		val	الملك	24.4.74	# 4 T T	210
				200		- -		•	205	3		t7= 1:	C_{1}	
Asp	Asn	Ser	Cys		Phe	G_Y	Leu			Arg	GTĀ	Vai	GTH	225
				215	_	m 3	~ 1		220			ጣጌ አንት	₽×'n	
Met	Arg	Phe	Thr		Pro	GLY	Phe	Pro		Ser	PLO	T Ä T	FTO	240
			٠	230		- 3	_	•	235	%	717	7 cm	Sar	
His	Ala	Arg	Cys		Trp	ALA	Leu	Arg		ASD	YTA	ASD		255
				245	•	~	Dia a		250	71.7	Sar	Care	Asm	
Leu	Ser	Leu	Thr		Arg	Ser	Pne	Asp	265	VIG	267	Cys	1100	270
		·	_	260	** 1	mb	t7- 7	Ch esa		- مرين -	T.all	Ser	Pro	
Arg	Gly	Ser	Asp		val	'InT	Val	TAT	280	4174	.neu	501		285
	_		- 3	275		C1 -	T	Cara		محارث	ጥ	Prò	Pro	
Glu	Pro	Hıs	Ala		Val	GLII	Leu	CYS	295	÷ ++	+ <u>7</u>		* * •	300
	_	_	æ³.	290 Dha	77.1 ~	C	Com	C1n		₹7 ⇒ 1	T.e.ii	T.eu	Tle	
Tyr	Asn	Leu	Thr		HIS	261	267	.4111	310	v a.r.	ПСС			315
_	~ J		Asn	305	· ~1	71	7	шiс		Gly	Phe	Glu	Ala	
Leu	TTE	Thr	ASN	320	· ·	WTA	ar d	IIT2	325	·		 -		330
Dla a	Db -	C1	Leu		7~~	Mat	Ser	Sar		Glv	Glv	Ara	Leu	
Pne	hue	GIII	•	335	wra	Mec	Der		340			~ <u>J</u>		345
T - 10	* 7		Gly		Dha	1 sn	Ser	Pro		Tvr	Pro	Gly	His	Tyr
гÃг	Ald	GTII	GTĀ	350	FILE	WOII			355	•		_		360
Dro	Pro	λεπ	Ile		Cvs	ሞከጉ	للبل	Asn			Val	Pro	Asn	Asn
LTO	ETO	WPII	115	365	Cys	1,744	D		370				٠	375
Gln	Hie	Val	Lys		Ser	Phe	Lvs	Phe	Phe	Tyr	Leu	Leu	Glu	Pro
الملد ت			10.J O	380	2 4				385	_				390
Glv	Val	Pro	Ala		Thr	Cys	Pro	Lys	Asp	Tyr	Val	Glu	Ile	Asn-
 1			-	395				-	400				•	405
Glv	Glu	Lvs	Tyr	Cvs	Gly	Glu	Arg	Ser	Gln	Phe	Val	Val	Thr	Ser
		— 	••	410					415					420
Asn	Ser	Asn	Lys	Ile	Thr	Val	Arg	Phe	His	Ser	Asp	Gln	Ser	Tyr
	. 2		-	425					430		-			435
Thr	Asp	Thr	Gly	Phe	Leu	Ala	Glu	Tyr	Leu	Ser	Tyr	Asp	Ser	Ser
	**		-	440				•	445		• .			450
Asp	Pro	Cvs	Pro	Gly	Gln	Phe	Thr	Cys	Arg	Thr	Gly	Arg	Cys	Ile
	*	-		455				_	460		•			465

	Arg	Lys	Glu	Leu		Cys	Asp	Gly	Trp	Ala	Asp	Cys	Thr	Asp	His 480
J	<u> </u>	3	Glu	Tan	470	Chres	Sar	ے برجی	Agn		Gly	His	Gln	Phe	
	Ser	Asp	ناط	ren	485	CVS	Ser.	Cys	, nam	490	<u> </u>		J	7 110	495
		T 775	Asn	Tire		Cvs	ivs	Pro	Leu		Tro	Val	Cys	Asp	
	CAR	m A D	warr	הלים	500	G	_	0		505				~	510
	Va 1	Asn	Asp	Cvs		Asp	Asn	Ser	Asp		Gln	Gly	Cys	Ser	Cys
	V 4.1	3	E	- 2	515					520		٠.			525
	Pro	Ala	Gln	Thr	Phe	Arg	Cys	Ser	Asn	Gly	Lys	Cys	Leu	Ser	Lys
					530					535					540
	Ser	Gln	Gln	Cys	Asn	Gly	Lys	Asp	Asp	Cys	Gly	Asp	Gly	Ser	Asp
					545					550					555
	Glu	Ala	Ser	Cys	Pro	Lys	Val	Asn	Val	Val	Thr	Cys	Thr	Lys	His
					560			-		565					-570
	Thr	Tyr	Arg	Cys	Leu	Asn	Gly	Leu	Cys	Leu	Ser	Lys	Gly		
					575		٠			580					585
	Glu	Cys	Asp	Gly	Lys	Glu	Asp	Cys	Ser		Gly	Ser	Asp		
					590					595					600
	Asp	Cys	Asp	Cys		Leu	Arg	Ser	Phe		Arg	Gin	Ala	Arg	
					605				• 7	610			CClaren	C1~	615
	Val	Gly	Gly	Thr		Ala	Asp	Glu	GTA		urp.	Pro	TIP	(141)	630
			1	- 7	620	~1	Q1-	~1	Wi a	625	Care	G117	Δla	Sar	•
	Ser	Leu	His			GTĀ	GTII	GTĀ	utz	640	CYS	GTA	A Sporter took		645
	T1 ~	Core	Pro		635	T 211	773]	Sar	Ala		His	Cvs	Tvr	Ile	
	TTG	per	LLO		650		VUL			655		4 2 -			660
	ASD	Ara	Gly				Ser	Asp	.Pro		Gln	Trp	Thr	Ala	Phe
	- v= F	3			665	·				670	-		-	-	675
	Leu	Gly	Leu	His	Asp	Gln	Ser	Gln	Arg	Ser	Ala	Pro	Gly	Val	Gln
		-			680					685					690
	Glu	Arg	Arg	Leu	Lys	Arg	Ile	Ile	Ser	His	Pro	Phe	Phe	Asn	Asp
					695	•				700				. •	705
	Phe	Thr	Phe	Asp	Tyr	Asp	Ile	Ala	Leu	Leu	Glu	Leu	Glu	Lys	Pro
		•			710					715					720
	Ala	Glu	Tyr	ser	Ser	Met	Val	Arg	Pro	Ile	Cys	Leu	Pro	Asp	Ala
					725	•				730					735
	Ser	His	Val	Phe	Pro	Ala	Gly	Lys	Ala		Trp	Val	Thr.	Gly	
					740	-				745					750

						•										
(Gly	His	Thr	Gl:	n Ty:	r Gly	Gly	/ Thr	: Gly	r Ala	a Lev	ı Ile	. Lei	ı Glı	ı Lys	5
					75	5				760					765	· ·
(Gly	Glu	Ile	Arg	y Vai	l Ile	Asr	ı Glr	Thr	Thr	Cys	Glu	Asr	ı Lei	ı Lev	1
					770					775	5				780)
]	Pro	Gln	Gln	Ile	e Thi	r Pro	Arg	, Met	Met	Cys	. Val	Gly	Phe	Lev	ı Ser	•
•					789	5	-	-		790)	•			795	
. (Gly	Gly	Val	Ası	Ser	: Cys	Gln	Gly	Asp	Ser	Gly	Gly	Pro	Leu	Ser	•
					800)				805					810	,
	Ser	Val	Glu	Ala	Asp	Gly	Arg	Ile	Phe	Gln	Ala	Gly	Val	Val	Ser	•
		•		•	815	5				820	,				825	
Ţ	rp	Gly	Asp	Gly		Ala	Gln	Arg	Asn			Gly	Val	Tyr		
					830					835					840	
A	\rg	Leu	Pro	Leu	Phe 845	Arg	Asp	Trp	Ile	Lys 850	Glu	Asn	Thr	Gly	Val 855	
•		•														
		<21			3	•				·						
			.1>		256							•				•
		<21 <21			PRT Unkr			•							•	
		<22			OTIVI	·			· .							
		<22			DOMA	TN										
•		<22					rote	ase	cata	lvti	c đo	main	of	heps	in (Heps)
		·				logo			. •	_				-	•	
-		<40	0>		3.			٠			•				•	
Δ.	*******	Tle '	Va I	Glv	Glv	Arg	Acn	ጥኮኍ	Sor	T.@11	Gly	Δ γ α	كتخيل	Pro	كشديل	
4 2.	- -	<u></u>	V CA.		5	***************************************	1100	***		10	·		L		15	
G.	ln '	Val:	Ser	Leu		Tyr	Asp	Glv	Ala		Leu	Cvs	'. Glv	Glv		•
					20					25				2	30	
Le	eu 1	Leu :	Ser	Gly	Asp	Trp	Val	Leu	Thr	Ala	Ala	His	Cys	Phe	Pro	
				-	35	_				40					45	
G.	lu A	Arg 1	Asn	Arg	Val	Leu	Ser	Arg	Trp	Arg	Val	Phe	Ala	Gly	Ala	
					50	•				55		•			60	
Vā	al A	Ala (Gln	Ala	Ser	Pro.	His	Gly	Leu	Gln	Leu	Gly	Val	Gln	Ala	
										7 0				:		
					65					70					75	
۷a	al V	/al 7				Gly	Tyr	Leu			Arg	Asp	Pro	Asn		
Vā	al V	/al 7				Gly	Tyr	Leu	Pro		Arg	Asp	Pro			
	,		ſyr	His	Gly 80	Gly Asp	. ·		Pro	Phe 85			•	· · ·	Ser 90	

```
Leu Pro Leu Thr Glu Tyr Ile Gln Pro Val Cys Leu Pro Ala Ala
                                     115
                                                          120
                 110
Gly Gln Ala Leu Val Asp Gly Lys Ile Cys Thr Val Thr Gly Trp
                                     130
                                                          135
                 125
Gly Asn Thr Gln Tyr Tyr Gly Gln Gln Ala Gly Val Leu Gln Glu
                                                          150
                                     145
                .140
Ala Arg Val Pro Ile Ile Ser Asn Asp Val Cys Asn Gly Ala Asp
                                                          165
                                     160
                 155
Phe Tyr Gly Asn Gln Ile Lys Pro Lys Met Phe Cys Ala Gly Tyr
                                     175
                                                          180
                 170
Pro Glu Gly Gly Ile Asp Ala Cys Gln Gly Asp Ser Gly Gly Pro
                                                          195
                                     190
                 185
Phe Val Cys Glu Asp Ser Ile Ser Arg Thr Pro Arg Trp Arg Leu
                                                          210
                                     205
                 200
Cys Gly Ile Val Ser Trp Gly Thr Gly Cys Ala Leu Ala Gln Lys
                                     220
                                                          225
                 215
Pro Gly Val Tyr Thr Lys Val Ser Asp Phe Arg Glu Trp Ile Phe
                                                          240
                                     235
                 230
Gln Ala Ile Lys Thr His Ser Glu Ala Ser Gly Met Val Thr Gln
                                                          255
                                     250
                 245
Leu
     <210>
      <211>
                225
      <212>
                PRT
     <213>
                Unknown
     <220>
     <221>
                DOMAIN
                Serine protease catalytic domain of Scce
     <223>
                homologous to similar domain in TADG-15.
     <400>
                4
Lys Ile Ile Asp Gly Ala Pro Cys Ala Arg Gly Ser His Pro Trp
                                                          15
                                     10
                 5
Gln Val Ala Leu Leu Ser Gly Asn Gln Leu His Cys Gly Gly Val
                                                          30
                                     25
                20
Leu Val Asn Glu Arg Trp Val Leu Thr Ala Ala His Cys Lys Met
                                                          45
                                     40
                35
```

Asr	ı Glı	ı Tyr	Thi		His	Leu	Gly	Ser		Thr	Leu	Gly	As <u>r</u>	Arg
3		.	_	50	_		_		55	-1		!	_	60
Arg	J Ala	a Gin	Arg		Lys	Ala	Ser	ŗŻZ		Phe	Arg	Hls	Pro	Gly
Th. +		- Mh -a	~1~	65 mb	***	77 1	7	3	70	34 - -	7	tr_ 1	-	75
τλr) ನಿರ್ವ	Thr	·	80	ura	val	ASN	Asp	ьец 85	Mec	Leu	val	гуs	
Asn	Ser	Gln	Ala		T. 	Sár	Ser	Mot	•	Tue	Tare	77=7	2 20	90
****		<u> </u>		95	عب ب	JC1		X 2 C	100			VCL	ar 9	105
Pro	Ser	Arg	Cys		Pro	Pro	Gly	Thr			Thr	Val	Ser	
,		-	-	110					115	-				120
Trp	Gly	Thr	Thr	Thr	Ser	Pro	Asp	Val	Thr	Phe	Pro	Ser	Asp	Leu
	,			125					130					135
Met	Cys	Val	Asp	Val	Lýs	Leu	Ile	Ser	Pro	Gln	Asp	Cys	Thr	Lys
				140					145		•		•	150
Val	Tyr	Lys	Asp	Leu	Leu	Glu	Asn	Ser	Met	Leu	Cys	Ala	Gly	Ile
				155			-		160.					165
Pro	Asp	Ser	Lys		Asn	Ala	Cys	Asn		Asp	Ser	Gly	Gly	Pro
-				170	_,				175	· •	_			180
<u>ren</u>	Val	Cys	Arg		'I'nr	Leu	GIn	GLY		Val	Ser	Trp	GLy	
Dha	Dro	Cys	ci.	185.	Dwa	7 ~~	7 ~~	Drea	190	*7 7		ſſſĥ~	~1~	195
	rio	Cys		200	PLO	ASII	ŸPħ		205	vai	TAT	7 777	GIII	210
Cys	Lvs	Phe			Tro	Ile .	Asn			Met	Lvs	Lvs	His	
	-			215	<u>-</u> -				220		_1 ~			225
	<21	0~	. c	5			·						;	
	<21			225										
	<21			PRT		•							-	
	<21			Jnkno	wn									
	<22	0>	-	•			-							
	<22	1>	· · I	OMAI	N					. •				
	<22	3>	S	Serin	e pr	otea	se c	atal	ytic	dom	ain	of.t	ryps	sin
			(Try)	hom	olog	ous	to s	imil	ar d	lomai	n in	TAL	XG-15.
	< 4.0	0>	5	;										
Lys	Ile	Val (Gly (Gly 7	Yr P	Asn (ys (Glu (Glu A	lsn S	Ser V	Jal :	Pro '	Tyr
		٠		5				1	1:0		٠ .	•		15
.Gln '	Val :	Ser I	Leu A	Asn S	Ser C	aly 1	yr i	His H	he C	ys (Gly (Gly s	Ser 1	Leu
			2	20				2	25				;	3 0

```
Ile Asn Glu Gln Trp Val Val Ser Ala Gly His Cys Tyr Lys Ser
                                                          45
                                     40
                35
Arg Ile Gln Val Arg Leu Gly Glu His Asn Ile Glu Val Leu Glu
                                                          60
                                     55
                50
Gly Asn Glu Gln Phe Ile Asn Ala Ala Lys Ile Ile Arg His Pro
                                                          75
                                     70
                65
Gln Tyr Asp Arg Lys Thr Leu Asn Asp Ile Met Leu Ile Lys
                                                          90
                                     85
                80
Leu Ser Ser Arg Ala Val Ile Asn Ala Arg Val Ser Thr Ile Ser
                                                          105
                                    100
                95
Leu Pro Thr Ala Pro Pro Ala Thr Gly Thr Lys Cys Leu Ile Ser
                                                          120.
                                     115
                110
Gly Trp Gly Asn Thr Ala Ser Ser Gly Ala Asp Tyr Pro Asp Glu
                                                          135
                                     130
                125
Leu Gln Cys Leu Asp Ala Pro Val Leu Ser Gln Ala Lys Cys Glu
                                                          150
                                     145
                140
Ala Ser Tyr Pro Gly Lys Ile Thr Ser Asn Met Phe Cys Val Gly
                                                          165
                                     160
                155
Phe Leu Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly
                                                          180
                                     175
                170
Pro Val Val Cys Asn Gly Gln Leu Gln Gly Val Val Ser Trp Gly
                                                          195
                                     190
                185
Asp Gly Cys Ala Gln Lys Asn Lys Pro Gly Val Tyr Thr Lys Val
                                                          210
                                     205
                200
Tyr Asn Tyr Val Lys Trp Ile Lys Asn Thr Ile Ala Ala Asn Ser
                                                          225
                                     220
                215
     <210>
                б
     <211>
                231
     <212>
                PRT
     <213>
                Unknown
     <220>
                DOMAIN
     <221>
                Serine protease catalytic domain of chymotrypsin
     <223>
                (Chymb) homologous to similar domain in TADG-15.
     <400>
                6
Arg Ile Val Asn Gly Glu Asp Ala Val Pro Gly Ser Trp Pro Trp
                                                          15
                                     10
```

Gln	Val	Ser	Leu	Gln	Asp	Lys	Thr	Gly	Phe	His	Phe	Cys		
				20		. •			25					. 30
Ser	Leu	Ile	Ser	Glu	Asp	Trp	Val	Val			Ala	His	Cys	Gly ·
				35		,			40					45
Val	Arg	Thr	Ser	Asp	Val	Val	Val	Ala		Glu	Phe	Asp	Gln	
				.50					55					60
Ser	Asp	Glu	Glu	Asn	Ile	Gln	Val	Leu		Ile	Ala	Lys	Val	-
			-	65				_	70			_	~ ?	75
Lys	Asn	Pro	Lys	Phe	Ser	Ile	Leu	Thr		Asn	Asn	Asp	Ile	
				80					85	_		ion1	T7- 1	90
Leu	Leu	Lys	Leu	Ala	Thr	Pro	Ala	Arg		Ser	GIn	'L'nr	Val	
·				95		_			100	 1		3.7		105
Ala	Val	Cys	Leu	Pro	Ser	Ala	Asp	Asp		Pne	Pro	Ala	GIÀ	
				110					115	7	70		71-	120
Leu	Cys	Ala	Thr	Thr	Gly	Trp	Gly	Lys		гЛZ	JÄI	ASN		
				125			•		130	-	D	Τ	•	135
Lys	Thr	Pro	Asp	Lys	. Leu	Gln	Gin	Ala		Leu	Pro	Leu	Leu	150
				140	_		_	<i>a</i> n 7 .	145	3	T7.		7	
Asn	Ala	Glu	Cys	Lys	Lys	Ser	Trp	GTA		Arg	TTE	1117	asp	165
				155			a 1	77- T	160	50*	ر م	Mot	Gla	
Met	Ile	Cys	Ala	Gly	Ala	Ser	GTĀ	vaı	175	per	CAP	Me c	GTA	180
_			_	170	¥7 7	C+		Tarm	-	G1 ₁₂	Δla	Tren	ጥኩ፦ -	
Ser	Gly	Gly	Pro	Leu	Agi	Cys	GTU	гЛг	190	GTĀ	ALA	115	444	195
v 3 1	0 1	77 7	** 1	185	(T)	~1·•	Ca r	7) CY		Cre	Car	Thr	Ser	
Val	GIŢ	TIE	Val	Ser	TID	GIY	267	ಶಾಗಿ	205	-	-			210
D	~1	۲7 ٦	(The exact	200 Ala	7~~	7727	- הלים	Twe	•	Tle	Pro	Tro	Val	
Pro	GTĀ	Val	тАт	215	wr â	Val	جديد بل خ		220			<u>F</u>		225
í ve	Tla	Len	7 J =	Ala	A en									
тÃъ	*T=	neu	ALG	230	11011					•				
	-2°	10>		7	-					·				
		11>		, 255										
•		1.2>		PRT										•
-		13>		: Unkn	OWII									
		20>												
		21>	-	DOMA	IN							•	· . ·	
		23>				rote	ase	cata	lyti	c do	main	of	fact	or 7
		-												ADG-15
					•		_							

						•								
	<4	0.0>		7	-									
Arg	Ile	Val-	Gly	Gly	Lys	Val	Cys	Pro	Lys	Gly	Glu	Cys	Pro	Trp
				5					10					15
Gln	Val	Leu	Leu	Leu	Val	Asn	Gly	Ala	Gln	Leu	Cys	Gly	Gly	Thr
			. '	20		·			25	•				30
Leu	Ile	Asn	Thr	Ile	Trp	Val	Val	Ser	Ala	Ala	His	Cys	Phe	Asp
				35					40					45
Lys	Ile	Lys	Asn	Trp	Arg	Asn	Leu	Ile	Ala	Val	Leu	Gly	Glu	His
				50					55		-			60
Asp	Leu	Ser	Gļu	Hiş	Asp	Gly	Asp	Glu	Gln	Ser	Arg	Arg	Val	Ala
_				65		-			70					75
Gln	Val	Ile	Ile	Pro	Ser	Thr	Tyr	Val	Pro	Gly	Thr	Thr	Asn	His
			-	80					85	-	,			90
Asp	Ile	Ala	Leu	Leu	Arg	Leu	His	Gln	Pro	Val	Val	Leu	Thr	Asp
				95			_		100			-	•	105
His	Val	Val	Pro	Leu	Cys	Leu	Pro	Glu	Arg	Thr	Phe	Ser	Glu	Arg
				110					115		•	-		120
Thr	Leu	Ala	Phe	Val	Arg	Phe	Ser	Leu	Val	Ser	Gly	Trp	Gly	Gln
				125					130				1	135
Leu	Leu	Asp	Arg	Gly	Ala	Thr	Ala	Leu	Glu	Leu.	Met	Val	Leu	Asn
	•	•		140					145	•				150
Val	Pro	Arg	Leu	Met	Thr	Gln	Asp	Cys	Leu	Gln	Gln	Ser	Arg	Lys
				155					160			-		165
Val	Gly	Asp	Ser	Pro	Asn	Ile	Thr	Glu	Tyr	Met	Phe	Cys	Ala	Gly
				170					175					1,80
Tyr	Ser	Asp	Gly	Ser	Lys	Asp	Ser	Cys	Lys	Gly	Asp	Ser	Gly	Gly
				185					190					195
Pro	His	Ala	Thr	His	Tyr	Arg	Gly	Thr	Trp	Tyr	Leu	Thr	Gly	Ile
				200					205			•		210
Val	Ser	Trp	Gly	Gln	Gly	Cys	Ala	Thr	Val	Gly	His	Phe	Gly	Val
	v			215					220					225
Tyr	Thr	Arg	Val	Ser	Gln	Tyr	Ile	Glu	Trp	Leu	Gln	Lys	Leu	Met
				230					235					240
Arg	Ser	Glu	Pro		Pro	Gly	Val	Leu	Leu	Arg	Ala	Pro	Phe	Pro
				245					250					255
	<23	10>		8										•
	<2.	11>		253								•		

			,	
<212>	PRT			
<213>	Unknown			
<220>				,
<221>	DOMAIN			.
<223>			Lytic domain	
		·	(Tpa) homo	logous to
	similar do	main in TAI)G-15.	
<400>	8			
Arg Ile Lys	Gly Gly Leu Ph	ne Ala Asp	Ile Ala Ser	His Pro Trp
	5		10	15
Gln Ala Ala	Ile Phe Ala Ly	s His Arg	Arg Ser Pro	Gly Glu Arg
	20		25	30
Phe Leu Cys	Gly Gly Ile Le	eu Ile Ser	Ser Cys Trp	Ile Leu Ser
• • •	35		40	45
Ala Ala His	Cys Phe Gln Gl	u Arg Phe	Pro Pro His	His Leu Thr
	50	•	55	60
Val Ile Leu	Gly Arg Thr Ty	r Arg Val	Val Pro Gly	Glu Glu Glu
	65		70	75
Gln Lys Phe	Glu Val Glu Ly	s Tyr Ile	Val His Lys	Glu Phe Asp
	80		85	90
Asp Asp Thr	Tyr Asp Asn As	sp Ile Ala	Leu Leu Gln	Leu Lys Ser
	95	•	100	105
Asp Ser Ser	Arg Cys Ala Gl	n Glu Ser	Sër Val Val	Arg Thr Val
•	110	·	115	120
Cys Leu Pro	Pro Ala Asp Le	eu Gln Leu	Pro Asp Trp	Thr Glu Cys
	125.	•	130	135
Glu Leu Ser	Gly Tyr Gly Ly	s His Glu	Ala Leu Ser	Pro Phe Tyr
•	140		145	150
Ser Glu Arg	Leu Lys Glu Al	a His Val	Arg Leu Tyr	Pro Ser Ser
	155		160	. 165
Arg Cys Thr	Ser Gln His Le	eu Leu Asn	Arg Thr Val	
	170	•	175	180
Met Leu Cys	Ala Gly Asp Th			
	185		190	195
Leu His Asp	Ala Cys Gln Gl			
<u>.</u>	200		205	210
•		•		

```
Leu Asn Asp Gly Arg Met Thr Leu Val Gly Ile Ile Ser Trp Gly
                                                          225
                  215
                                      220
 Leu Gly Cys Gly Gln Lys Asp Val Pro Gly Val Tyr Thr Lys Val
                  230
                                      235
                                                          240
 Thr Asn Tyr Leu Asp Trp Ile Arg Asp Asn Met Arg Pro
                                      250
      <210>
                 9
      ,<211>
                 2900
      <212>
                 DNA
                Homo sapiens
      <213>
      <220>
      <223>
                 SNC19 mRNA sequence (U20428)
      <400>
                                                           50
 cgctgggtgg tgctggcagc cgtgctgatc ggcctcctct tggtcttgct
                                                          100
 ggggatcggc ttcctggtgt ggcatttgca gtaccgggac gtgcgtgtcc
 agaaggtett caatggetac atgaggatea caaatgagaa ttttgtggat
                                                          150
                                                          200
gectacgaga actecaacte caetgagttt gtaageetgg ceageaaggt
                                                          250
gaaggacgcg ctgaagctgc tgtacagcgg agtcccattc ctgggcccct
                                                          300
 accacaagga gtcggctgtg acggccttca gcgagggcag cgtcatcgcc
                                                          350
tactactggt ctgagttcag catcccgcag cacctggttg aggaggccga
gcgcgtcatg gccaggagcg cgtagtcatg ctgcccccgc gggcgcgctc
                                                          400
                                                         450
cctgaagtcc tttgtggtca cctcagtggt ggctttcccc acggactcca
aaacagtaca gaggacccag gacaacagct gcagctttgg cctgcacgcc
                                                          500
gcggtgtgga gctgatgcgc ttcaccacgc cggcttccct gacagcccct
                                                          600
accccgctca tgcccgctgc cagtgggctg cggggacgcg acgcagtgct
                                                         650
gagetacteg agetgaeteg cagettgaet gegeetegae gagegeggea
                                                         700
gcgacctggt gacgtgtaca acaccctgag ccccatggag ccccacgcct
                                                         750
ggtgagtgtg tggcacctac cctccctcct acaacctgac cttccactcc
                                                         800
ctcccacgaa cgtcctgctc atcacactga taaccaacac tgacgcggca
                                                         850
tcccggcttt gaggccacct tcttccagct gcctaggatg agcagctgtg
                                                         900
gaggccgett acgtaaagcc caggggacat tcaacagccc ctactaccca
ggccactacc cacccaacat tgactgcaca tggaaaattg aggtgcccaa
                                                        950
caaccagcat gtgaaggtgc gcttcaaatt cttctacctg ctggagcccg 1000
gcgtgcctgc gggcacctgc cccaaggact acgtggagat caatggggag 1050
aaatactgcg gagagggtc ccagttcgtc gtcaccagca acagcaacaa 1100
gatcacagtt cgcttccact cagatcagtc ctacaccgac accggcttct 1150
tagctgaata cctctcctac gactccagtg acccatgccc ggggcagttc 1200
acgtgccgca cggggcggtg tatccggaag gagctgcgct gtgatggctg 1250
```

	ggcgactgca	ccgaccacag	cgatgagcto	aactgcagtt	: gcgacgccgg	1300
	ccaccagttc	acgtgcaaga	gcaagttcto	, caagetette	: tgggtctgcg	1350
	acagtgtgaa	cgagtgcgga	gacaacagcg	acgagcaggg	, ttgcatttgt	1400
	ccggacccag	accttcaggt	gttccaatgg	gaagtgeete	tcgaaaagcc	1450
	agcagtgcaa	tgggaaggac	gactgtgggg	acgggtccga	cgaggcctcc	1500
	tgccccaagg	tgaacgtcgt	cacttgtacc	aaacacacct	accgctgcct	1550
	caatgggctc	tgcttgagca	agggcaaccc	tgagtgtgac	gggaaggagg	1600
	actgtagcga	cggctcagat	gagaaggact	gcgactgtgg	gctgcggtca	1650
	ttcacgagac	aggctcgtgt	tgttgggggc	acggatgcgg	atgagggcga	1700
	gtggccctgg	caggtaagcc	tgcatgctct	gggccagggc	cacatctgcg	1750
	gtgcttccct	catctctccc	aactggctgg	tctctgccgc	acactgctac	1800
	atcgatgaca	gaggattcag	gtactcagac	cccacgcagg	acggccttcc	1850
	tgggcttgca	cgaccagagc	cagcgcaggc	cctggggtgc	aggagcgcag	1900
	gctcaagcgc	atcatctccc	accccttctt	caatgacttc	accttcgact	1950
	atgacatcgc	gctgctggag	ctggagaaac	cggcagagta	cagctccatg	2000
	gtgcggccca	tctgcctgcc	ggacgcctgc	catgtcttcc	ctgccggcaa	2050
-	ggccatctgg	gtcacgggct	ggggacacac	ccagtatgga	ggcactggcg	2100
1	cgctgatcct	gcàaaagggt	gagateegeg	tcatcaacca	gaccacctgc	2150
,	gagaacctcc	tgccgcagca	gatcacgccg	cgcatgatgt	gcgtgggctt	2200
(cctcagcggc	ggcgtggact	cctgccaggg	tgattccggg	ggacccctgt	2250
(ccagcgtgga	ggcggatggg	cggatcttcc	aggccggtgt	ggtgagctgg	2300
9	ggagacgctg	cgctcagagg	aacaagccag	gcgtgtacac	aaggctccct	2350
(ctgtttcggg	aatggatcaa	agagaacact	ggggtatagg	gaccadaacc	2400
ć	acccaaatgt	gtacacctgc	ggggccaccc	atcgtccacc	ccagtgtgca	2450
	gcctgcagg	ctggagactc	gcgcaccgtg	acctgcacca	gcgccccaga	2500
ć	acatacactg	tgaactcatc	tccaggctca	aatctgctag	aaaacctctc	2550
Ç	gcttcctcag	cctccaaagt	ggagctggga	gggtagaagg	ggaggaacac	2600
t	ggtggttct	actgacccaa	ctggggcaag	gtttgaagca	cagctccggc	2650
2	agcccaagtg	ggcgaggacg	cgtttgtgca	tactgccctg	ctctatacac	2700
9	gaagacctg	gatctctagt	gagtgtgact	gccggatctg	gctgtggtcc	2750
t	tggccacgc	ttcttgagga	agcccaggct	cggaggaccc	tggaaaacag	2800
				ctcccaggtg tttcttttta		
	<210>	10			•	
	<211>	20		•	· .	•
	<212>	DNA				
	<213>	Artific	ial Sequenc	ce,		
	<220>					

```
<221>
                 primer
       <223>
                 Forward primer for analysis of overexpression
                 of TADG-15 mRNA by quantitative PCR.
      <400>
                 10
 atgacagagg attcaggtac
                                     20
      <210>
                 11
                 20
      <211>
      <212>
                 DNA
      <213>
                 Artificial Sequence
      <220>
      <221>
                 primer
                 Reverse primer for analysis of overexpression
      <223>
                 of TADG-15 mRNA by quantitative PCR.
      <400>
                 11
                                   20
gaaggtgaag tcattgaaga
      <210>
                 12
      <211>
                 17 .
      <212>
                DNA
      <213>
                Artificial Sequence
      <220>
      <221>
                primer
     <223>
                Forward primer for analysis of \beta-tubulin mRNA
                expression by quantitative PCR.
     <400>
                12
tgcattgaca acgaggc
                                 17
     <210>
                13
     <211>
                17
     <212>
                DNA
     <213>
                Artificial Sequence
     <220>
     <221>
                primer
     <223>
                Reverse primer for analysis of _-tubulin mRNA
                expression by quantitative PCR.
     <400>
                13
ctgtcttgac attgttg
                                17
```

:.**:**.

	<	210>		14										
	<	211>		242									-	
	<	212>		PRT	1									
	<	213>		Hom	o sa	pien	s `							
	<	220>												
	<	221>		DOM	AIN									
	. <	223>		Ser	ine]	prot	ease	cat	alyt	ic d	omai	n of	TAD	G-15.
٠.	<	400>		14										
Arc	y Va.	l Val	L Gly	g Gly	Thr	Asp	Ala	Asp	Glu	Gly	Glu	ı Try	p Pro	Trp
	-	•		5	-				10					15
Glr	ı Val	l Ser	: Leu	His	Ala	Leu	Gly	Gln	Gly	His	Ile	Cys	Gly	/ Ala
				20					25					30
Ser	Lei	ı Ile	Ser	Pro	Asn	. Trp	Leu	Val	Ser	Ala	. Ala	His	Cys	Tyr
				35					40					45
Ile	Asp) Asp	Arg	Gly	Phe	Arg	Tyr	Ser	Asp	Pro	Thr	Glr	Trp	Thr
				. 50					55			•		60
Ala	Phe	Leu	Gly	Leu	His	Asp	Gln	Ser	Gln	Arg	Ser	Ala	Pro	Gly
				65					70					75
Val	Gln	Glu	Arg	Arg	Leu	Lys	Arg	Ile	Ile	Ser	His	Pro	Phe	Phe
				80					85					90
Asn	Asp	Phe	Thr	Phe	Asp	Tyr	Asp	Ile	Ala	Leu	Leu	Glu	Leu	Glu
				95					100					105
Lys	Pro	Ala	Glu	Tyr	Ser	Ser	Met	Val	Arg	Pro	Ile	Cys	Leu	Pro
				110					115					120
Asp	Ala	Ser	His	Val	Phe	Pro	Ala	Gly	Lys	Ala	Ile	Trp	Val	Thr
				125		·			,130		•			135
Gly	Trp	Gly	His	Thr	Gln	Tyr	Gly	Gly	Thr	Gly	Ala	Leu	Ile	Leu
				140		,			145				2	150
Gln	Lys	Gly	Glu		Arg	Val	Ile	Asn		Thr	Thr	Cys	Glu	
_	_		_	155					160					165
Leu	Leu	Pro	Gln		Ile	Thr	Pro	Arg		Met	Cys	Val	Gly	Phe
				170					175				••	180
Leu	Ser	Gly	GLY		Asp	Ser	Cys			Asp	Ser	Gly	Gly	
_		,		185					190					195
Leu	Ser	Ser			Ala	Asp	Gly .				Gln	Ala	Gly	
		_		200					205	•	•			210
val	Ser	Trp			Gly	Cys .	Ala			Asn	Lys	Pro	Gly	
				215					220					225

PCT/US99/03436

WO 99/42120

Tyr Thr Arg Leu Pro Leu Phe Arg Asp Trp Ile Lys Glu Asn Thr
230 235 240

Gly Val

INTERNATIONAL SEARCH REPORT

International application No. PCT/US99/03436

A. CLASSIFICATION OF SUBJECT MATTER IPC(6) :Please See Extra Sheet.										
US CL: 530/324; 536/23.5; 435/320.1, 69.1, 6 According to International Patent Classification (IPC) or to both national classification and IPC										
B. FIE	LDS SEARCHED									
Minimum	documentation searched (classification system follo	wed by classification symbols)	·							
U.S. :	530/324; 536/23.5; 435/320.1, 69.1, 6									
Document	ation searched other than minimum documentation to	the extent that such documents are included	d in the fields searched							
Electronic	data base consulted during the international search	(name of data base and, where oracticable	scarch terms used)							
DIALO		(. Journal terms dised)							
C. DOC	CUMENTS CONSIDERED TO BE RELEVANT		,							
Category*	Citation of document, with indication, where a	appropriate, of the relevant passages	Relevant to claim No.							
Y,P	TANIMOTO, H. et al. Cloning and Expression of TADG-15, A 1-11 Novel Serine Protease Expressed in Ovarian Cancer. Proceedings of the American Association for Cancer Research. March 1998, Vol. 39, page 648, especially page 648.									
Y,P	O'BRIEN, T.J. et al. Cloning and Expression of TADG-15, A I-11 Novel Serine Protease Expressed in Ovarian Cancer" Tumor Biology. August 1998, Vol. 19, Supplement No. 2, pages 33, especially page 33.									
	·									
	·									
Furthe	er documents are listed in the continuation of Box (C. See patent family annex.								
Spe	eral categories of cited documents:	"T" later document published after the inter								
'A' doc to b	ument defining the general state of the art which is not considered of particular relevance	the principle or theory underlying the								
L* dos	ier document published on or after the international filing date untent which may throw doubts on priority claimts) or which is	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone								
zbec	d to establish the publication date of another citation or other teason (as specified)	"Y" document of particular relevance; the	claimed invention cannot be							
O* document referring to an oral disclosure, use, exhibition or other combined with one or more other such documents, such combination means										
document published prior to the international filing date but later than "" document member of the same patent family the priority date examed										
	pate of the actual completion of the international search. Date of mailing of the international search report.									
26 APRIL	1999	19MAY 1999								
lame and m. Commissione Box PCT Washington.	ailing address of the ISA/US or of Patents and Trademarks	Authorized officer YVONNE EYLER ALLICE LICE								
acsimile No		Telephone No. (703) 308 0196								

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER:

International application No. PCT/US99/03436

IPC (6):
A61K 38/00; C07K 5/00, 7/00, 16/00, 17/00; C07H 21/04; C12N 15/00, 15/09, 15/63, 15/70, 15/74; C12P 21/06; C12Q

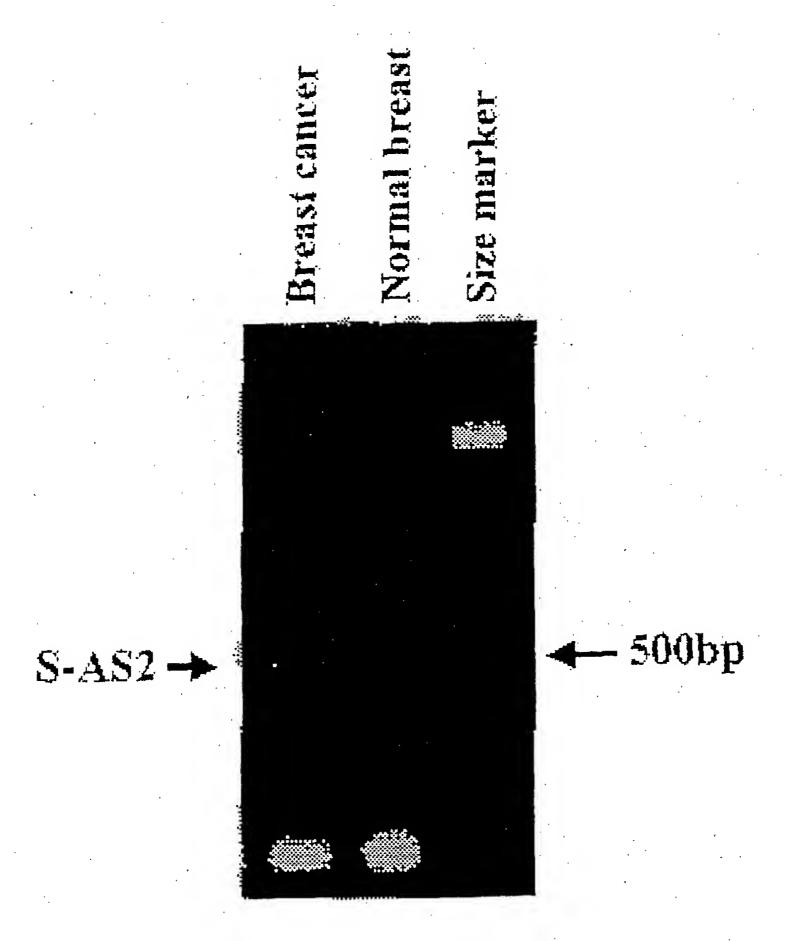
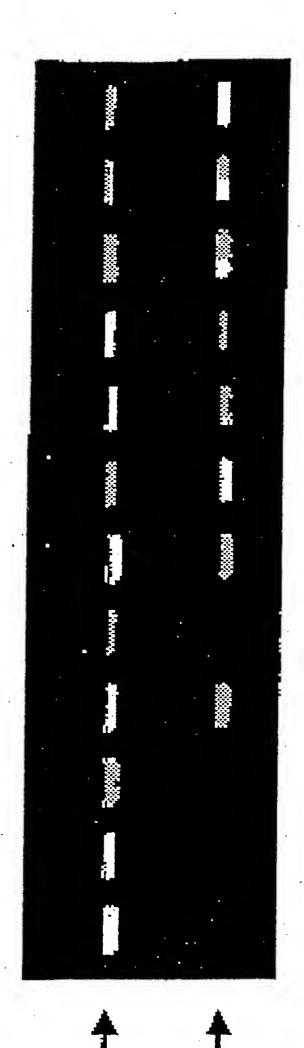


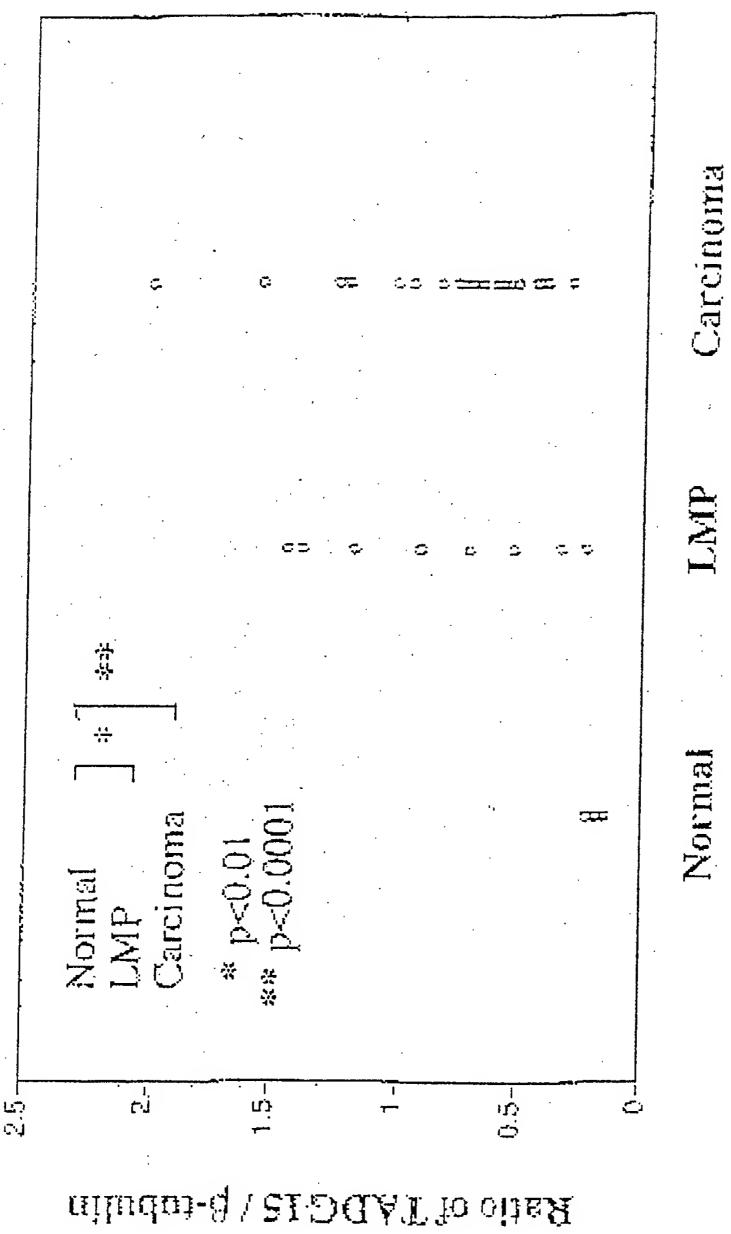
FIG. 1

																		r
VAÇASPHGLC HDOSORSAPC	OTLG. DR.F HMIEVLEG.P	li II	TTRVVPGE	NTTTYGOOT!	TTSPDVTFI	KIKIMANKTI	QLLIMGALBI KHEALSPEYI		MRTCELVSM			-						-
LSRWAVFAGA	KARIQVRLGE KSRIQVRLGE	RISOVVVAGE	Tw1LGR	GKICTVIGNG GKAIWVIGNG	GTTCTVSGOG	GTLCATTGMG	WRESLYSGWG WTECELSGYG		CEDSISHTPR	SSVEADGR		COKDEA	CLMDGR	Heps Heps	ન	Try	Foo 7	
FE RNRV	T	GV	QERFFPHHL.	ASHVEPA	RCE THE	Addoda	KTESEKTLAE LPPADLQLPD	÷	_	COGDSGGPL.		CKGDSGGF	CGGCSGGFLV	MO: 3)	i i	NO:	1 PC:	
DWVLTABACE			CMILSBARCE	EXIQEVCLES		OTVSAVCLES	OESS VVRIVC		5IDA	5VDS	6808	GVÖG SKDS	-	CEÇ. ID		SEQ.	an tend In	
HLCGGSLLSC HIOGESLISC	H. CGGVLYNE CGGSLINE	HFCGGSLISE OLCGGILINT	FLCGGILISS	HLSS, FLFLT ELEK, PREYS	KINS OBBLS KICS BENTH	KLAT. PARFS	KLHU-FVVLT QLKSDSSRCA				MECWSFLEG	VMICAGAS YMFCAGYSDG	NMLCAGDIRS	SEASGNUTOL				PKFGVLLKAF
AYDG.A	LSGNQL NSGYHF	QOKTGE	arhrrspcer *	EENSNDIALV D. YDIALL	HANDENIA T		CITMHDISLD DNDIALL		FYGW. OIKP		YPG. KITS	WGR. RITD KVGDSPNITE	LLWRTVTD	EWIFORIKTH	KWINDTMKKH	KWIKWIIAAN	PWVCKILABA	EWLOR LMKSE OWIRDMHRF~
GRMPWQVSL. Gedomoner	GEMPNOVAL. NSVPYQVSL.	GSWPWOVEL. GECPWOVLL.	ashpnqaaif	GYLPERDPNS		KWFKFSILT.	VHKEFDODIY		ISNDYCNGAD	INCTICE. N	LSONKCERS.	LSNAECKKS. MTODCLOOSR	YPSSRCTSOR	GVYTEVSDER	GVYTRUPLER	GVYTKVYNYV	GVYBEWTKLI	GVYTRVSQYI GVYTRVTNYL
RIVGGROTSL	KIIDGAPCAR KIVGGYNCEE	RIVNGEDRUP RIVGGKVCFK	RIKGGLFADI	LGWQAWYHG	RORIKASKSE	IOVLKIAKVE	USKKVAUVII EOKFEVEKYI		GVLOEARVEL	LILOKGEIRV	DELQCIDARY	DKLOOAALFL ELMVLMVFRL	ERLKEAHVEL	T. GCALAOKP	D.GCRORNKE	D. GCAOKHKP	SDACS, TSSP	Q.GCATVGHF .LGCGQKDVP
								•										

normal ovary normal ovary m LMP s LMP s LMP s LMP s LMP s carcinoma s carcinoma s carcinoma s carcinoma s carcinoma s carcinoma



<u>교</u>



SUBSTITUTE SHEET (RULE 28)

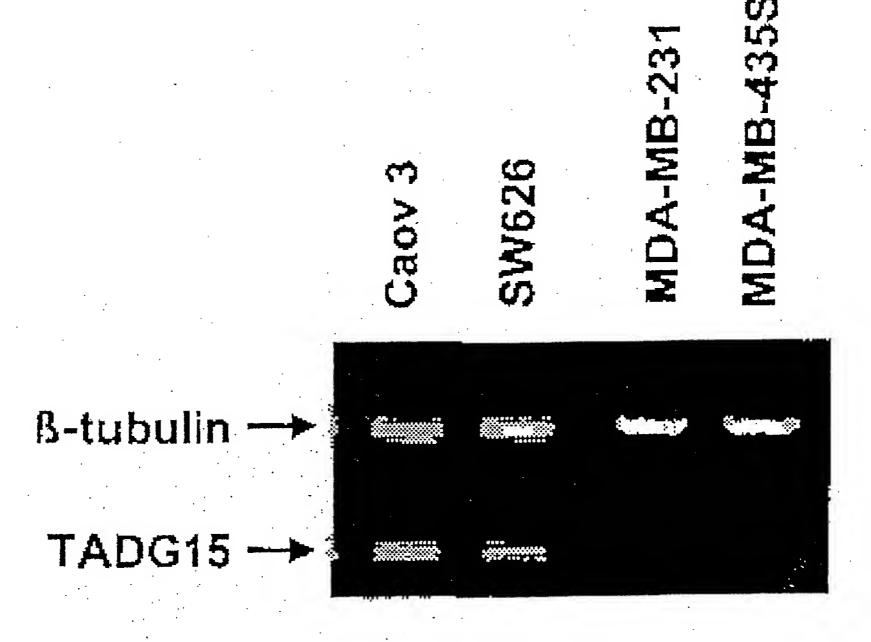


FIG. 5

Ovarian cancer

Breast cancer

Colon cancer Prostate cancer Lung cancer

ß-tubulin

TADG15 →

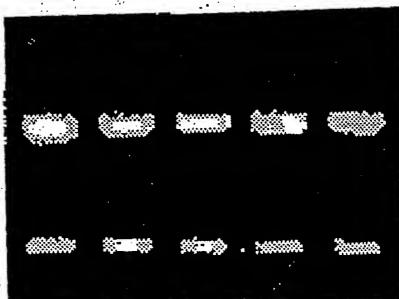


FIG. 6

WO 99/42120

Pancreas

Skeletal Muscle

P.B. Leukocyte *

andastri llom2

Kidney

Liver

Bunj

Ricin

Неац

Colon

Ovary

Testes

Prostate

sumyhT

Spieen

Kiquek

Liver

กีนกา

niona

Clear Cell Carcinoma

Muclinous Carcinoma

Serous Carcinoma

Normal Overy

Endometriold Carcinoma

Placenta

3147

ORF (23 - 2587)

1993 2428 Original subclone

Clone A

2150

Clone B

129

Clone C

.

OFFICE CHEST (DIES 28)

	TO ALCHEOUS CONTROL OF THE RESIDENCE OF
_	TO ACTUAL AND A LEGISLATION OF A REPORT OF
1	THE RESIDENCE OF THE RE
111	COLLECTION CONTROL OF THE PROPERTY AND
	一条一条一样,我们们在一条一条一条一样,这一样,这一样,这一个一个,我们就是这个人的人,我们就没有这个人的人,我们就会会会会会会会会会会会会会会会会会会会会会会
213	CITETYSCISIONAL TOSCHARGE CONTRACTOR CONTRAC
• • •	THE TRANSPORT OF THE PROPERTY OF THE FOREST PROPERTY PROPERTY OF THE FOREST PROPERTY OF THE FOREST PROPERTY PROP
127	THE PROPERTY AND A CONTRACTOR OF THE PROPERTY OF THE RESERVE OF TH
: E4E	,我们就是一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个
	CONTROL TO THE PROPERTY OF THE
†14.T	CONTROL OF THE PARK TO A STATE OF THE PARK TO THE PARK
670	CACCTERGYOGTECCTITCCCERROSSRETCENNIACRATING GROWN CECTOSTATION CONTINUENCE CON
	一种,我们就们就没有一个。" "我们就是一个,我们就是一个,我们就是一个,我们就是一个,我们就是一个,我们就是一个,我们就是一个,我们就是一个,我们就是一个,
123	contribute the property of the
	CONCRETE STATEMENT OF THE STATEMENT OF T
Su' l	TOTAL TRANSPORTED TO A STATE OF THE PARTY OF
116	CHARLEST OF THE FOREST OF THE
٠ بر	The state of the s
77:	CATCACACTUATAACCACACTCAGCCOCCOCCOCCOCCOCCACACTCACACACTCACACACTCACACTCACACACTCACACACTCACACACTCACACACTCA
	- 4
1941	THE THE THE PROPERTY OF THE PR
	The state of the s
1771	THE TRACE CONTROL OF THE PROPERTY OF THE PROPE
	A G T G S K D S Y E 1 8 W TO THE THE PROPERTY OF THE PROPERTY
13.53	THE HOLD STATE OF THE PROPERTY
	,是一种一种一种一种一种一种一种一种一种一种一种一种一种一种一种一种一种一种一种
	GREENARDOG CARDOGETTETTAGETANNIGENETIE IN CHECK S. B. F. C. P. G. P. F. C. A. F. L. A. A. P. C. D. A. D. P. C. P. G. P. F. C. A. F. C. A. P. L. F. A. V. C. D. A. V. P. P. C. A. P. K. F. C. M. P. C. D. A. V. P. P. C. A. P. K. F. C. M. P. C. D. A. G. H. D. P. T. C. A. P. K. F. C. M. P. V. C. D. A. V. P. P. C. A. P. K. F. C. M. P. C. D. A. G. H. D. P. T. C. A. P. K. F. C. M. P. V. C. D. A. G. H. D. P. T. C. A. P. K. F. C. M. P. P. C. M. P. C. M. P. P. C. M. P. P. C. M. P. M. P. C. M. P. P. C. M. P. P. C. M. P. M. P. P. C. M. P.
_ 1 :	
7553	
1981	THE CONTROL OF THE STATE OF THE
IEV.	TO THE REPORT OF A B L L S P H H L V E A A (E) C V L D D R C F R P B D P L C H C H C H C H C H C H C H C H C H C
	THE THE RESERVE THE PARTY OF THE PROPERTY OF THE PROPERTY OF THE PARTY
114	to t
•	1 1 0 0 E 1 C G A H L 1 S F A STATE OF THE S
LLDE	TOTAL PROPERTY OF A STATE OF A ST
ፈ ሳ ውሳ	GENTICETEGRATION TO CONTROLL TO THE REST OF THE REST O
. .	THE REPORT OF TH
374.1	CONTRACTOR OF THE FORM THE FOR
135 1	CONTRACTOR OF THE PROPERTY OF THE PARTY OF T
2422	CENTRALISMENT CONTROL OF THE PARTY OF THE PA
	CENTRALISMENT CONTROLLED CONTROLL
1577	CETAMORRACIONALITACIO
***	HANDACCROCCTOURGETTOURCETTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTC
1211	THE STREET WITH STREET WITH STREET ST
#16 €.÷	# 8 P G V Y T R L P G P P G P G P G P G P G P G P G P G
7:07	CONSTITUTION OF THE PROPERTY OF THE MOST OF THE MOST OF THE PROPERTY OF THE PR
22 6 2	The transfer of the second of
	TENTIFICATION OF EACH TO NO. 1)
37.7%	,我们还是这个大学的,我们就是一个大学的,我们就是一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个一个

L. 1 : Karaks Consumses Segmente

O: Conserved among actide of extelytic total H.D.S

Transpearment dimery

Figure 9.

1	MGSDRARKGG	GGPRDFGACL	KYMSRHEXVM	GLEEGVESLP	VNNVKKVEKE	
51	GPGFWVVLMA	VLIGLLLVLL	GIGFLVWHLQ	YRDVRVQKVF	NGYME I THEN	7
101	FYDAYENSINS	TEFVSLASKV	KDALKLLYEG	VPFLGPYHKE	SAVTAFSEGS	<i>34</i>
151	VIAYYWSEFS	IPORLVEEAE	RVMASERVVM	LPPRARSLKS	EVVTEVVARP	•
201	TOSKTVORTO	DNSČSEGLAA	RGVELMRFT'T	PGFPDSFYPA	HARČOMALRG	
251	DADSVLSLTF	RSFDLASÕDE	RGSDLVTVYN	TLSPMEPHAL	VQLČGTYPPS	
301	импъневой	VLLITLITAT	Errhpgfeat	FFQIARMSSČ	GGRLRKAQGT	3
351	FNSPYYPGHY	PPNI DČTWN (EVPNNQHVKV	SFKFFYLLEP	GVPAGTČEKO	
401	YVEINGEKYÖ	GERSOFVVTS	NSMK1TVREH	SDQSYIDTGF	LAEYLSYDSS	
951	DECEGGETCR	TGRCIRKELR	COGWADCTON	SDELNCSCDA	GROFTCKNKE	
501	CKPLFWVCDS	VNDCGDHSDE	QGCSCPAQTF	ROSNGKOLSK	SQQCNGKDDC	4
551	GOGS DEPASCE	KVNVVTCTKE	TYRCLNGLOL	GKCMPECDGK	EDCSDGSDEK	•
601	DCDCGLRSFT	ROARVVGGTD	ADEGEW PWQV	STAALGQGHI	CGASLISPNW	
651 .	LVSAABITYID	DRGFRYSOPT	QWTAFIGLHD	QSQRSAPGVQ	ERRLKRIISH	
701	PERNORTEDY	OIVITEFEKE	afyssmyrfi	CLPDASHVER	AGKAIWVTGW	5
751	GHTQYGGTGA	LILOKGEIRV	INOTTCENLL	POQITPRAMO	VGFLSGGVDS	Į.
801	CQGR\$GGPLS	SVEADGRIFO	àGVV\$WGDG0	AQRNKPGVYT	RLPLERDWIK	
851	ENTGY (SEQ.	ID NO: 2)				

: Conserved cysteine residue

: Possible N-linked glycosylation site

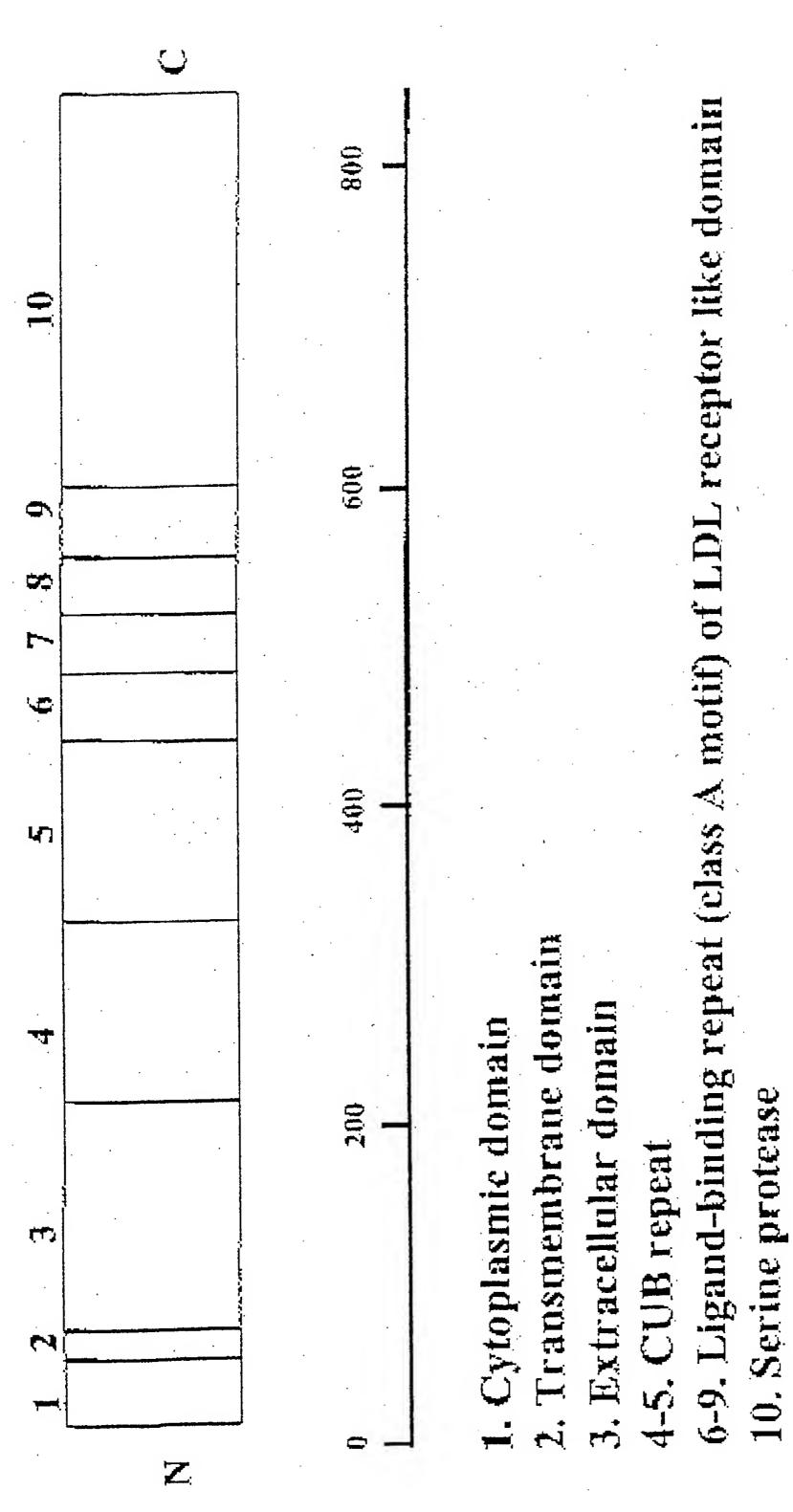
SDE : Conserved SDE motif

: Potential cleavage site

: Conserved amino acids of catalytic triad H, D, S

- 1. Cytoplasmic domain
- 2. Transmembrane domain
- 3. CUB repeat
- 4. Ligand-binding repeat (class A motif) of LDL receptor like domain
- 5. Serine protease

FIG. 10



12/13

	MITES STON STONE STONE STONE STANDED CACAMBEN ABTERNATE SOTHING TOTAL COSSOLL FLITTINGS TITLE SOURCE TOTAL CONTROLS TITLE SOURCE	human. Heme supiens Enkaryonar; mitoche Vertebsatz; Enthari 1 (Mases) to 7' Theng.S., Cai,X., (Enkly game in Heme Unpoblished 2 (Doses) to 7' Theng.S. Hiract Subsission Subsited (Mi-Jeff) Endlesi Lad (Mi-Jeff)	sequence. Indital subaryotes: (Ly Prinstes: Galary SOD(SOD(SAPINO SAPINO MANGEDOU, Show Show (MANGEDOU, SIGN(), Show MANGEDOU, SIGN(),	nami maminada ng.l. and ibi naminada namina hapabi	inchiana in of China		
	T	eneds: Topasaseessee	्तृहरकृष्यम्भयोगेषे तर्गकरम् २००० इ.स.च्याच्या	LTOSA (TYCTOSA)	(1110mereestanning)	eerstatesseriument	93
		(C) 9:,					This
82 0	Bastaclaetti	CARCALOSCO ACODAS	icategaerereecstik.				
. 1			anggerer er seller i seller		······································		1 01
1 ([11]]]]]]]] [20][20][20][]]]		CCTCCTCTTEGTCTT \$276	PRINCIPAL CONTROLL	S S TETROCATT SECALT.	accuração do de autorio	140
1 101 A			erterikastitaseelis 1111-1111 1110 1111 1111 24 terereriyastigiseelis				200
983 6 1 201 6	ARGENCECECTI 	eregregrerietekeekek Hilligietekeekekeekeekeekeekeekeekeekeekeekeeke	MCCCATTOCTEESCCIIIA {{{{ 1	::::::::::::::::::::::::::::::::::::::	070193624461011111 07019362446607	CARROLLAND AND AND AND AND AND AND AND AND AND	30 0
482 5 1 201 7	actrotorer Hillitii Actrotorer	LITTORICATORICANIO LITTORICANIONICENSIONI LITTORICANIONICENSIONI	ioggipaterracies de la composition della composi	I FLETCRISCECC (117 17 17 17 4 CFCSTCRISCECC.A	SIR BOGERTASTONIGE LITTI I I I I I I I I I I I I I SINGOGOSTASTONIGE	1905050505050505050505050505050505050505	561 399
1 460 C	11111111111 የፕሮመየጵዱ።የመጀመ	Tateatcaccicaedere			<u>i hodajski padadite</u>	Charttegoldtoniches.	498
। ধুপুট ক	11111111111111 	i i i i i i i i i i i i i i i i i i i	CONSTRUCTOR CONTRACTOR	TANTATE FICH I	indificies seel.	Toeneneach, com	597
593 G	i i i i i i i i i i i i i i i i i i i		TITEACCTTGEGTCGT3T2 	ACERECECESSIANCE	inggreen all a seco	<u>CLACACCTEMSCECCHE</u>	69.6
1 697 G	illilililii i Argototototo	in die	CONCOCTODOTOTRUS 	Coloratorio	CONTROCERCES COTTO	THE TENTE VENEZULATION	763
ነ የፅብ ፍ	(11 1113]] Taratronen, Jo	<u>;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;</u>	OCCAPITATE OCCAPITATE CONTRACTORS	: '' ಸಾರ್ಥವಾಗಿ ಬಿಡಿಯ	TSTEERMOODETTREE	Transproperiodelect	##1
\$97 C		i i i i ji	CEMENTYGEET GEFORTS THILLIH HILLIH FOR EGGEOTTGGET GEFORE	erecket excepted	Charles and the contract of th	Mágaggettermise.	561
1375 S 6 982 T	odiadoregios 111111111111111111111111111111111111	gerroepograditais 	landergebookkagestike: 	27502522702476 	S MAGRARIA TECESRA 	HEVERACED STACES OF THE STATE O	780C 7874
l Teage	11111111111111111111111111111111111111	THE HEALTH CHAINSTEE STATES	TTOURCTENSATURGEOUS 	kaloom, ahaqeet	TTITTTAGETELETE	TICTACCACTOCADIA	1161
1 23 53 ¢	i i i i i i i i i i i i i i i i i i i		Becognesia (Colored Colored Co	25131.41.45.25.25.29	attani, escerció	independent and a second	1250
1478 A 1281 A	ACTECAETTECO 111111111111111111111111111111111111	regergessericherster Hillight (IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	gyockadkadvacitatii 	######################################	SECTORESEES OF SELECTION OF SECTION OF SECTI	aceastateteareacanea Alii IIIIIIIIIIIIIIII Aceastateteareacanea	1354
						•	

]575)778	. CERACELEGAGESTECRATICPOCCA. COCKENCETTERATORIZACIONALEGENASTORIZACIONALEGAGACASTELU TEGRARGIOCIALITORI 1675 - (111111111111111111111111111111111111
ነፅ ^ግ ተ	COCACESSTIC GROUPSECCTECTSCOCCARGET CARGET CATE TO THE TREE TO A COCACEST CONTRACT TO THE TERM CARGET ASSESSMENT OF THE STATE OF THE ST
3774	CECTERGIETEREGGENAGERGEREGENEELEGENEELEREGENEELERETERGRECHNEGGEGGEGGENTERTEREGLEATREGGTTTTTTCGG 1813
157 4 1874	CONTRACTORECONALIZATION OF THE CONTRACTOR OF THE
)673	TITE ENGTOXIANOCITATEDESTERNING CONTRACTORIST AND CONTRACTORIST CONTRACT
(779	TESTCTSTSSSSAGTESTADITEATEACAGAGAGAGATCASSTACTCAGAGCOCAGGGAGTESAGGGGCCTTCTSSGCTTGCAGGGAGAGGGAGGCAGGCAGGCAGGCAGGCAGGC
2074	Chececooffesterrorserscorescoreraterrorserscoreraterrorsers (ACTICACTICACTICACTICACTICACTICACTICACTI
2174	CARANCESCARASTACACTERATACIONECECECATOTARIO CORRESCOCICECATOTETTATUTECCISCOCECARSOCIATOTORISTANCOCOCTICO PARA
2274	GRONANCOSCADA CARETO DA GOTO DE GOLDANO DE LO CARRETO DE CARRETO D
3334	CACRORSOLATATERAGRAMICALINGUE DE LO PERSAGRAMA PROPERTAGRAMICA DE LO CONTRACTOR DE LO CACADO DE LA VI CACECUECECATURA DE L'ALCONOMICA DE L'ACOMO DE L'ACECUE DE L'ACECUE DE L'ACECUECE DE L'ACECUECE DE L'ACECUE
2171	CACINGGGGATGTGGGTGGGTTGGTGAALAGDAGTTGGGTGGGTGGGTGGGTGTGTGGGGGGGTGGGGGGGG
3234	ANCTICCASECCEGTETESTERESTAGGEREELE. CETEOGUTCHERGEREELERECCHEGUSTETAGMERLESETECCTETETTTEGESMANGERTCHAG 2302
2273	AGRACACTOSSIGNATARIOGRACIOS DE CARACTOS DE CARACTOS CONTROLAS CONTROLAS DE LA CONTROLA CONTROLA CARACTOS CONTROLAS DE LA CONTROLA CARACTOS CONTROLA CARACTOS CONTROLAS DE LA CARACTOS DE LA CARACTOS CONTROLAS DE LA CARACTOS CONTROLAS DE LA CARACTOS DEL CARACTOS DE LA CARACTOS DEL CARACTOS DE LA CARACTOS DEL CARACTOS DE LA CARACTOS DEL CARACTOS DE LA CARACTOS DEL CARACTOS DEL CARACTOS DE LA CARACTOS DE LA CARACTOS DEL CARACTOS DE LA CARACTOS DE LA CARACTOS DE LA CARACTOS DEL CARACTOS DE LA CARACTOS DE LA CARACTOS DEL CARACTOS DE
2573	EMACESTERSTGERSCHOSSSCHOOMERINGERSTERRANTSTEURSSESTERRANTSTEINERARASTTOTSGETTGETCRESSTERRANTSTEINERSTEINER 2770 1 711
2772	ABTGENSCHEDEN, BOTH SPRIGGEGAGG, ACHOT DE VEGT TOTAGTER DOCKET VIGES STAARGT TOTAGENIAGE SCOTET COCCETAMED CENASO (\$6.68)
2664	AGTGENGTTGTGAGGTAGAGGGGAGGGAGGGAGGGAGAGAGAGAG
785 <i>3</i>	TORGERARDICATE TORGETA CTORE, ETRETETAS MEMORGANIZADES
40.52	TERS TINGENTALING TO THE TERMINATE OF THE PROPERTY OF THE PROP
30£9 2823	CHECK TORTHOGOGRAPHIC TOTAL AND CONTRACT CONTRAC

Figure 12 (cont.)

SEQUENCE LISTING

<110>		en, Timothy oto, Mirotos		•	
<120×		.5; An Extra			
<130>	D6064P	CT			٠
<140>	FCT/US	99/03436			
<141>	1999-0	2-18			
<150>	បទ 09/	627,337			÷
<151>	1998-0	2-20			
<160>	14				
<210>	1				
<211>	3147				
<212>	LMA				
<213>	\mathcal{H} çmo s	apiens			
<220>			•		
<223>	cIMA a	equence of	TADG-15		
<400>	3				
tcaagagcee	cotoggggta	ccatggggag	cgatcgggcc	ರಧಾಷಕ್ಷುಗೆದಿದ್ದ	50
ಿಕಡೆದೆಡೆಡೆದಿಂದ	gaaggzöttö	<u>ರಿದ್ದಿದ್ದರೆರೆದಿತ್ತಾ</u>	tcaegtaces	ರ್ ರದದದ್ದಾರ್ವತ್ತ	100
gagaaagtga	atggcttgga	<u>ತಿಡಿಕಳಿಗೆ ಹೇಗಿ</u>	gagtteetge	cagtoaacaa	150
cgtcaagaag	gtggaaaagc	atggcccggg	gagatgggtg	gtgatggaag	200
ccgtgctgat	oggactests	ttggtcttgc	tggggàtogg	etteetagte	250
tggcatttgc	agtaccygga	egtçeçtgte	cagaaggtot	tosatggota	300 .
catgaggatc	acasstyaya	attttgtgga	tgcctacgag	aactccaact	35Û
ccactgagtt	tgtaagcctg	<u>ರಂಭಾರ್ಥವಾಕ್</u> ತರೆದೆ	tgaaggacgc	gctgaagctg	400
ctgtacagog	gagtoccatt	cetadaceee	taccacaagg	aqteggetgt	450
gacggccttc	ಇ ದೇವಿಚಕ್ತಿಡಿದಿ	gegteatege	ctactactqq	totgagttca	500
gcatcccgca	goacctggtg	gaggaggccg	agegogteat	ਰੋਧੋਵਵਧੇਤਾਵੋ	550
cgcgtagtca	tgotgoddaa	gegggegeige	tocctgaagt	cetttytygt	600
cacctoagtg	gtggetttee	ccacgga c tc	ceaaacagta	cagaggadco	650
<u> ಇದಿದೆ ಇದ್ದ ಇದೆ</u>	ctgcagettt	ggcctgcacg	cccacaatat	ggagetgatg	700 ,
egetteacca	agazaggatt	ccctgacago	ccctaccccg	cccatgcccg	750
ctgccagtgg	geactgeggg	ನನೆಡೆತ್ತುಬಿಂದಬಿತ	ctcagtgctg	agecteacet	800
	tgaccttgcg			•	850
acggtgtaca	acaccctgag	coccatadad	ceceaegeee	tggtgcagtt	900
	taccetccct				
acgtootgct	cateadacts	a taa ddaaqa	ctdaacaaca	gcatcccgge	1000

				•	
tttgaggtca	cottottota	getgeetagg	atgagcagct	<u>atagaagaaq</u>	1050
cttacgtaaa	<u>ಡಿಂಂಂತರಿಕೆಕೆಕೆ</u>	cattcaacag	cocctactac	ccaegonact	1100
RCCCACCCAG	cattgactgc	acatogaaca	ttgaggtgco	CESCAROCAG	1150
catgtgaagg	tgagottcaa	attottetac	ctgctggage	paggagtgad	1200
tgcgggcacc	tgccccaagg	actacgtgga	gatcaatggg	gagaaatact	1250
dedavadad	gtoccagtto	gtogtoacca	çcaacagcaa	caagatoaca	1300
gttogottoo	actoagatoa	gtoctacaco	gacaccygct	tottagotga	1350
atacctctcc	tacgactoca	gtgacccatg	೧ ೮೮೪೪೪೪	ttdsogtgod	1400
descâdadaca	gtgtatccgg	aaggagetge	gotytyatyy	dtgggdogad	1450
tgcaccgacc	adagogatija	gotosacigo	agttgcgacg	ರರಭಿಭೇರನಾರದ	1500
gttcacgtgc	aagaacaagt	totgcaaged	catettatgg	gtatgagaca	1550
gtqtqaacqa	ctgcygagac	ಷಷ ಂ ಪಿದ್ದಂದೆತ್ತಾಗೆ	agcaggggtg	cagttgtccg	1600
gcccagacct	traggtytto	caatgggaag	typototoga	esegocay c s	1650
gtgcaatggg	asggacgact	årdddaciii	gtccgacgag	gaataatgaa.	1700
ccaeggtgas	cgtcgtcact	tgtaccaaac	acacctaccg	ctdcctcast	1750
gggatatgat	tgagcaayyy	caaccetgag	tgtgacqqqqa	aggaggactg	1800
tagogaeggc	tcagatgaga		ctgtgggctg		1850
cgagacaggc	togtgttgtt	ដំដីជីជិជិជិជិជនបដ់ដ	atgoggatga	वैवैवैटवेश्वेद्वे	1500
cestagasaga	taagcotgon	tgatatggga	cagggccaca		1950
ttecctcatc	totoccaact	ggetegtete	tgccgcacac	tgotadatog	2000
atgacagagg	attoaggtac		·	adcottocta	2050
ggottgcacy		de de såedes		aggagcgcag	2100
gatamagaga			caatgacttc		
atgacatcgc	gotectages	ctggagaaac	cggcagagta	cagetecatg	2200
gtg c ggccca	tetgeetgee	•	catgictics		
ggccatctgg	gtcacgggct		ccagtatgga		
cgctgatcct	gcaasagggt		tcatcaacca		
gagascotco	tgeegeagea		cycatgatgt		2400
<u>ಂದ ಕಿರಡಿದ್ದರೆದ್ದರ</u> ದ			tgattccggg		2450
			aggccggtgt		2500
<u> </u>	dodctcadad	gaacaagcca	ggcqtqtaca	caaggotoou	2220
			tggggtatag		2000
			catogtocac		2650
			ctycaccago		
			caastctgcc	_	2750
			ggaggtagaa		
ctggtggttc	tactgaecca	actgggggca	aaggitigaa	gachongoot	2000
ccccccccaa	coccaagety	ggccgaggcg	ogtttgtgta	tatotgcoto	لائا لي ش

·			
	දේශ්ෂ්තීයක්ට අද්ධිමේක්ට්ට්රි		
aggtyytygg sc	tgooggat otgggotg	rtg gggcesttgg	geoacgetet 3000
೬ರ್ಷಾಗಿ ಕಾರ್ಥಿಕ ಕಾರ್ಥಿಕ	aggetogga ggacectç	da sascadacad	gtatgagaat 3050
	scozgatac coggataga		
aaatgggtaa sa	acaatttat ttctttt	BBBBBBBBBBB CS	aaaaaa 3147
<210>	2		
<211>	855		•
<212>	PRT		
<213>	Homo sapiens	•	
<220>	Transfer and description	ುವರ್ಣ ೧೯ TATG-	15 encoded by cDNA.
<223>	2		
₹400 ≯		•	
Met Gly Ser A	asp Arg Ala Arg Ly	•	
	5	30	15
Phe Gly Ala G	Gly Leu Lys Tyr As		
	20	25	30
Gly Leu Glu G	alu Gly Wal Glu Ph		Asn Asn val Lys
·	35	40	•
Lys Val Glu I	ys His Gly Pro Gl		Agt ten std wra
	50	55	,
Val Leu Ile G	Sly Leu Leu Leu Va	•	Tie dry rue nen
	65 83	70 . - Ral ama Ral	
Ast LLb HTE r	en Gln Tyr Arg As	. 85 P AGT WIG AGT	90
	80 Met Arg Ile Thr As	•	
wer oth the r	95 The Tit wo	100	105
The how the B	een Ser Thr Glu Ph	7-7-7	
and went ser w	110	i15	120
Tare Revo bla I	len Lys Leu Leu Ty		Pro Phe Leu Gly
TYP TABLE FALLS	125	130	135
From Them His Y	ys Glu Ser Ala Va		Ser Glu Gly Ser
	140	145	150
Val Ile Ala T	hr Tyr Trp Ser Gl	u Phe Ser Ile	Pro Gln His Leu
	155	1.60	1.55
Val Glu Glu A	ala Glu Arg Val Me	t Ala Glu Glu	Arg Val Val Met
	170	175	130

Lau	Fro	Pro	Arg		Arg	Ser	Leu	Lyg	Ser 190	Fhe	Wal	Val	Thr	Ser
Val	Val	Als	Phe	185 Pro	Thr	Asp	5er	Γλέ	Thr	Val	Gln	arg	Thr	Gln
Asp	Asn	Ser	Cys	200 Ser		Gly	Leu	His	205 Ala	Arg	Gly	Val	Glu	210 Leu
				215				÷	220					225
Met	Arg	Fhe	Thr	Thr 230	Pro	Gly	Phe	Pro	Asp 235	Sex	Pro	Tyr	Fro	240
Ris	Alā	Arg	()/s		Trp	Ala	Leu	Ary	Gly	Asp	āla	Asp	Ser	Val
•			٠.	245					~ ~ ~			_		255
Leu	Ser	Leu	Thr		Arg	Ser	Phe	ASP	Leu 265	Ala	Sex	Cys	Asp	G1u 270
איזירי ב	Gly	Sar	ASD	260 Leu	Val	Thr	Val	Tyr	Agn	Thr	Leu	Ser	Fro	
1 ~~ "	· · · · · · · · · · · · · · · · · · ·			275				_	280					285
Glu	Pro	His	Als	Len	Val	Gln	Leu	Cys	Gly	Thr	Tyr.	Pro	Pro	
		•		290			-		295				T	300
TYI	<u>R≘n</u>	L≞u	Thr		His	Ser	Ser	Gln	Asn	Val	Γ÷ά	Leu	TIE	315
_		%	_	305	r ^{rq} 7 a.	7	س. س. الا	ш	310	C 11-	T ha		Ale	_
⊤ <i>9π</i>	11E	TET	ASII	320	다내	w.A	*****	UT S	Pro 325	ά+λ	£ 112			330
Phe	Phe	Gln	Γ≐π	-	Arg	Met	Ser	Ser	Cys	Gly	Gly	क्रिय	Pën	Arg
				335					340					345
Lys	Ala	Gln	Gly	Thr	Phe	Aen	Ser	Pro	$\mathbf{T}_{\mathbf{M}^{\mathrm{T}}}$	Tyx	Pro	Gly	His	
·				350					355					360
Pro	Pro	Asn	Ile		Cha	Thr	Trp	Asn	Ilė	Glu	Val	Fro	AEII	
				365	.	-1	.		370	rrh es	. I .=33	T = 11	C 127	375 Pro
Gln	His	Val		yai 380	Ser	Fne	TÀR	rne	Fhe 385	TVI	TICEN	Ή¢: rit	المراجد الم	390
(Z) ae	Men II	Pro		•	Tha	Cva	Pro	Lvs	Asp	Tyr	Val	Glu	Ile	
	4 54.4			395		,	,	,	400	_				405
Gly	Glu	Lys	Tyr	Cys	Gly	Glu	Arg	Ser	Gln	Phe	Val	Val	Thr	Ser
				410	-				415		·	-		420
Aen	Ser	Asn	Lys	Ile	Thr	Val	Arg	Fhe	His	Ser	Asp	Gln	Ser	Tyr
				425	-				430					435
The	qeA	Thr	Gly	•	Len	Ala	Glu	Tyr	Leu	Ser	Tyr	Asp		
_			_	440	, ,= 1		CH.)	/T.	445	o rd	መስፈ	Ti wakuta		.450 .ar
ASD	Fro	Cys	Pro		GIU.	Phe	Thr	CAE	Arg 460	THE	ĖΤΛ	មកភ្ន	ω y S	465
				455					لزادي به					

Arg	Lys	Glu	L∉u		Cya	rep	Gly	Trp		Asp	Çyş	<u>arte</u>	gzś	
				470					475	- 2	'	2. 2.	N.	.480
Ser	पुटर्स	Glu	Leu	Asn	Cys	Ser	DAR	ASP		GTA	HIS	Gin	rne	
				485					430		_		_	4.95
Cyrs	Lys	Ran	Lya	Phe	Cys	Lys	Fro	Leu	Phe	Trp	Val	ुर्ड	ASP	
				500					505					510
Wal	Asn	Asp	Cys	Gly	Asp	Asn	Ser	Asp	Giu	Gln	Gly	CAS	Ser	
				\$15			-		520					525
Pro	Ala	Gln	Thr	Phe	Arg	Cys	Ser	Agn	Gly	TAS	Cya	Ten	Ser	
				530					535		•			540
Sex	متلق	Glr	Çys	Asn	Gly	Lys	Asp	Asp	Cys	Çİy	Asp	Gly	Ser	gaf
				545					550					555
Glu	Ala	Ser	Cyd	Pro	Lys	Val	Asn	Val	Val	Thr	Cys	Thr	Lуз	His
				560					555					570
Thr	Tyr	Arg	Cys	Leu	Aşn	Gly	Leu	Cys	Leu	Ser	TAE	Gly	Asn	Fro
	• • .			575	- :				580		· .	. •		585
Glu	Cya	Asp	Gly	Lys	G1u	Asp	ுக	Ser	Asp	G1y	Ser	Asp	Glu	Lys
				590					595					600
Asp	Cys	Asp	Cys	Gly	Leu	Arg	Ser	Ph⊜	Thr	Arg	Glů	Ala	yrq	Val
				605		•			£10					£15
Wal	Gly	Gly	Thr	Asp	Ala	<i>kep</i>	Glu	G1y	Glu.	(£xP	Pro	Trp	Gln	Val
				620					625					630
Ser	Leu	Hís	Ala	Leu	Gly	Gln	Gly	Hi≊	Ile	Суз	Gly	Ala	Ser	Γŧή
				535					540	•				645
I1e	Ser	Pro	Asn	Trp	Leu	Val	Ser,	Ala	Ala	HįZ	<u>್ರ</u> ೀಕ	Tyr	Ilė	yen
				550					555					550
Asp	Ary	Gly	Phe	Arg	Tyrr	Ser	Asp	Pro	Thr	Gln	Trp	Thr	Ala	Fhe
		-		665					670					675
Leu	Gly	Leu	His	Asp	Gln	Ser	Glm	Arg	Ser	Ala	FID	Gly	Val	Gln
				ឥន្ធប្			•		585					690
Glu	Arg	Arg	Leu	Lys	Arg	Ile	Ile	Ser	His	Pro	Phe	Fhe	Asn	Asp
				695					700	•	-			705
Phe	Thr	Phe	qaA	Tyr	App	Ile	Ala	Leu	Leu	Glu	Leu	Glu	Lys	Pro
				710					715		,			720
Ala	Glu	Tyr	Ser	Ser	Met	val	Arg	ਮੁਸ਼ਨ -	IJ≢	Cys	Leu	Pro	Asp	Als
				725	•		•		730			,		735
Sar	Hi≎	Val	Phe	Pro	Ala	Gly	Lys	Ala	·Ile	Trp	Val	Thr	Gly	TIP
•				740					745	•				750

WO 99/42120 PCT/US99/03436

```
Gly His Thr Gln Tyr Gly Gly Thr Gly Ala Leu Ile Leu Glm Lys
                                                           755
                 755
                                      750
Gly Glu Ile Arg Val Ile Asn Gln Thr Thr Cys Glu Asn Leu Leu
                                     775
                                                          750
                 770
Pro Gln Gln Ile Thr Pro Arg Met Met Cys Val Gly Phe Leu Ser
                                                          795
                                      790
                 785
Gly Gly Val-Asp Ser Cys Glin Gly Asp Ser Gly Gly Pro Leu Ser
                                      805
                                                           810
                 300
Ser Val Glu Ala Asp Gly Arg Ile Phe Gln Ala Gly Val Val Ser
                                                          825
                                      820
                 815
Trp Gly Asp Gly Cys Ala Gln Arg Asn Lys Fro Gly Val Tyr Thr
                                      835
                                                          840
                 830
Arg Leu Pro Leu Phe Arg Asp Tro Ile Lys Glu Asn Thr Gly Val
                                                          B-5.5
                                      850
                 845
      <210>
                256
     <211>
     <212>
                PET
      <213>
                Unicropym
     <220>
     <221>
                DOMAIN
                Serine protease catalytic domain of hepsin (Heps)
      <223>
                homologous to similar domain in TADG-15
      <400>
                3
Arg Ile Val Gly Gly Arg Asp Thr Ser Leu Gly Arg Trp Pro Trp
                                                          15
                                     10
Glm Val Ser Leu Arg Tyr Asp Gly Ala His Leu Cys Gly Gly Ser
                                                          30
                                     25 -
                 20
Leu Leu Ser Gly Asp Trp Val Leu Thr Ala Ala His Cys Phe Pro
                                                          45
                                     40
                 35
Glu Arg Asn Arg Val Leu Ser Arg Trp Arg Val Phe Ala Gly Ala
                                                        5.0
                                     55
                50
Val Ala Gln Ala Ser Pro His Gly Leu Gln Leu Gly Val Gln Ala
                                                          75
                                     70
                 65
Val Val Tyr His Gly Gly Tyr Leu Pro Phe Arg Asp Pro Asm Ser
                                                          90
                                     85
                 80
Glu Glu Asn Ser Asn Asp Ile Ala Leu Val His Leu Ser Ser Pro
                                                          105
                                     100
                 95
```

Leu	Pro	Leu	Thr		Tyr	Ilė	Gln	Fro	val 115	Суз	Leu	Pro	Ala	Ala 120
~~~.	;÷1 vs	Als	T 2011	110	a mon	G15x	Tare	T1=		Thr	Val	Thr	Gly	
لآجادا	in Tii	WTIZ	₩.	125	******	"A" who 플립	ma n	~~~	130		•		-	135
GTV	Asn	Thr	Gln		Tyr	Gly	Gln	Gln	BLA	Gly	Val	Leu	Gln	Glu
ر _ا مده سه	2 6 6 7 4	23.22		140	-	_			145					150
Ala	Arg	Val	Pro	Ila	Ile	ser	Asn	Asp	Val	Cys	Asn	Gly	Als	Asp
				155					160					165
Fhe	Tyr	Gly	Asn	Gln	Ile	lly's	Pro	Lys		Fhe	Cy's	Ala	Gly	
				170					175	u	<b></b>	ord	con.	180
Fro	Glu	Gly	Gly		Asp	Ala	Cys	Gln		Asp	Sar	tà T.P.	GTY	195
		Cys		185	. <del></del>	<b>ተ</b> ነ ኤ	E ^M paymen	g v.a	190 mbr	ው _ሞ ል	ይተነተተ	<b>ጥ</b> ተንነት ን	Athir	
Phe	Val.	Cys	Gin	200	DOI.	TT=	SOT	1.24 H	205		****	~-~		210
The state	ralar.	Ile	Wal	• •	Thro	Gly	Titat	Gly		Ala	Leu	ăla	Gln	Lye
in ji w	OT.	بة طييف	y 54.11	215	20,5			-	220					225
Pro	Glv	Val	Tyr	Thr	Lys	Val	Ser	Asp	Pha	Arg	Glu	Trp	Ile	Phe
<b>4</b> ,			_	230					235			,		240
Ģln	Ala	Ile	Lys	Thr	His	Ser	Glu	Ala	Ser	Gly	Wet	Val	Thr	Gln
		-		245					250					355
Leu		•												
	<23	10>		4										
	<b>&lt;2</b> )	11>		225						•				
	- 4 <b>2</b> :	12>		PRT				٠		-				
	_	13>		Unkn	own					,				
		20>		ም-ሜኔ <i>ር</i> ን	T) 7									
		21.> 23>		DOMA Cari		rate	ase	caba	lvti	උ ප්ප	main	of	Seca	
	~ <u>~</u>	44 eff -5%					o si							
	<41	)O>		4										
Tare	<b>T</b> ] =	Ile	Jr. errein	glv.	ង់គ្រ	Fra	Övs	Ala	Arg	Glv	Ser	His	Pro	Trp
، ت ازلد	TTG	<b>T</b> 'T <u>C</u>		.5	474-44				10					15
Gln	Val	Ala		•	Ser	Gly	Asn	Gln	Leu	His	Cys	Gly	Gly	.Val
-				20					25					30
Leu	val.	).sn	Glu	Аrц	TIP	Val	Leu	Thr	Ala	Aln	His	CAS	Lys	Met
				35				-	40					45

PCT/US99/03436

													_	
Aan	Glu	Tyr	Thr	Val	His	Leu	Gly	Ser	Asp	Thr	Γ≅л	GIV	AED	
				50					5.5					50
Ary	Ala	Gln	Arg	Ils	Γλέ	Als	Ser	The	Sax	Phe	Yrd	His	Fro	
				65					70			_	_	7 <b>5</b>
Tyr	Ser	Thr	Gln	Thr	His	Val	Asn	Asp	Leu	Met	Léu	Val	Гув	
				<b>\$</b> 0					8.5					<del>9</del> 0
Asn	Ser	Gln	Ala	Arg	Leu	Ser	Ser	Met	Val	Lys	Lys	Val	Arg	
				95					100			1	<b>~</b>	105
Pro	Sex	Arg	Сув	Glu	Pro	Pro	Gly	Thi	Thr	CAS	Thr	Val	Ser	
				110					115				•	120
TTP	Gly	Thr	Thr	Thr	Ser	Pro	<b>As</b> ₽		Thr	Phe	Pro	Ser	velt	
				125			•		130		_	<b>6</b> %	<b>673</b> ha a a	135
Met	Cys	val	ARD	Val	Lys	Leu	Ile	Ser	Pro	Gin	ASP	Cys		
				140					145	4.		w. 1		150
Val	$T_{\mathcal{F}}$ r	Γλε	asp	Leu	Leu	Glu	Asn	Sex	Met	hen	ျှာ	WTG	GT.A.	
		٠		155		_		_	160	•	. <del></del>	መግኒል	~7	165
Pro	Asp	Ser	TÀR		Asn	Ala	Cha	Asn	Gly	ASD	ser	PTA	LT.A.	180
				170					175	**- 7	Cons	<u>Претта</u>	GT 18	
Leu	Val	Çys	Arg		Thr	Leu	GIR	G4V	Leu		ವಿ <b>ದ್</b> ತ	الريخ ببالد بال	ra T.Ž	195
				185		_		<del></del>	190		M - 14	المرائد الم	CI Text	
Phe	Fro	CAR	Gly		Pro	Agn	Aep	Fre	Gly	AST	тÄт	T T7T	77.77	210
			, <u>.</u>	200		<b>~</b> 7.	3	724.0	205	7.1 19	Lago	Tare	म् स्टब्स	-
Çyb	Lys	Phe	Thr		Trp	113	ASII	дын.	Thr 220	ner	ny a	n) a		225
				215			٠		قرا عند شد					
	<20	10>	·	5				-		•				
	<23	11:-		225			·							
	<20	12>		PRT						•				
	<21	13>		Unkn	OWIL									·
	423	20>												
	₹22	21>		DOMA							•	r		!
•		23>							lyti					
				(Try	) ho	molo	ಗ್ರಿಯಾಕ	to	siml	lar	gons.	in i	n in	DG-15.
	<4(	000		<b>5</b> .										
Lys	Ile	Val	Gly	GLY	$\mathcal{I}_{\mathcal{A}^{\mathcal{M}}}$	Asn	Cys	Glu	Glu	Asn	Ser	Val		Tyr ·
				5	•				ip					·15
Gln	Val	Ser	Leu	Agn	Ser	G1y	Tyr	His	Phe	Cys	Gly	Gly	Ser	
•				20					25					30

```
Ile Asn Glu Gln Trp Val Val Ser Ala Gly His Cys Tyr Lys Ser
                                                           45
                  35
                                      40
 Arg Ile Gln Val Arg Leu Gly Glu His Asn Ile Glu Val Leu Glu
                                                           60
                                      55
                  50
 Gly Asn Glu Gln Phe Ile Asn Ala Ala Lys Ile Ile Arg His Pro
                                                           75
                                      70
                  65
 Gin Tyr Asp Arg Lys Thr Leu Asn Asn Asp Ile Wet Leu Ile Lys
                                                           90
                                      35
                  巴口
 Leu Ser Ser Arg Ala Val Ile Asn Ala Arg Val Ser Thr Ile Ser
                                                           105
                                      100
                 95
 Leu Pro Thr Ala Pro Pro Ala Thr Gly Thr Lys Cys Leu Ile Ser
                                                           120
                                      115
                  110
 Gly Trp Gly Asn Thr Ala Ser Ser Gly Ala Asp Tyr Pro Asp Glu
                                                           135
                                      130
                 125
 Leu Gln Cys Leu Asp Als Pro Val Leu Ser Gin Ala Lys Cys Glu
                                                           150
                                      145
                 140
. Ala Ser Tyr Fro Gly Lys Ile Thr Ser Asn Met Phe Cys Val Gly
                                                           155
                 155
                                      190
 Phe Leu Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly
                                                           180
                                      175
                 170
 Pro Val Vai Cys Asn Gly Gln Leu Gln Gly Val Val Ser Trp Gly
                                                           195
                 185
                                      190
 Asp Gly Cys Ala Gin Lys Asn Lys Pro Gly Val Tyr Thr Lys Val
                                                           210
                                      205
                 200
 Tyr Asn Tyr Val Lys Trp Ile Lys Asn Thr Ile Ala Ala Asn Ser
                                                           225
                                      220
                 215
      <210>
                 б.
                 231
      <211>
      <212>
                 FRT
                 Unknown
      <213>
      <220>
      <221>
                 DOMAIN
                 Serine protesse catalytic domain of chymotrypsin
      <223>
                 (Chymb) homologous to similar domain in TADG-15.
      <400>
                 Ę
 Arg Ile Val Asn Gly Glu Asp Ala Val Pro Gly Ser Trp Pro Trp
                                                           15
                                      10
```

Gln	Val	Ser	Leu	Gln	Asp	Lys	Thr	Gly	Phe	His	Pha	Cys	GIY	Gly
				20					25					30
Ser	Leu	I1e	Ser	Glu	Asp	Try	Val	Val	Thr	Ala	Ala	His	CAa	${\tt Gl}_{{ extstyle Y}}$
				35					40			·		有量
val	Arg	TLT	Ser	Asp	val	Val	Wal	Pla	Gly	Clu	Fire	<u> Asp</u>	Gln	Gly
				50					55					50
Ser	Asp	Glu	Glu	Asn	Ile	Gln	Val	Leu	Lye	Ile	Ala	The	Wal	Ph <del>e</del>
				65					70					75
Lys	Asn	Pro	Lys	Fhe	Ser	Ile	Leu	Lur	Val	Agn	Aşn	Asp	Ile	Thr
				80					85					90
Leu	Lau	Lye	Γ≅n	Ala	Tha	Pro	als	Arg	Phe	Ser	Gln	Thr	V=1	Ser
				95					100		-			105
Ala	Val	Cys	Leu	FTO	Ser	Ala	Asp	Asp	asp	Phe	Pro	Ala	Gly	Thr
	•			110					115					120
Leu	Cys	Ala	Thr	Thr	Gly	TIP	Gly	Lys	Thr	Lys	Tyx	Aen	Ala	Asn
				125	•				130			•		135
Lys	Thr	Pro	Aep	Lys	Leu	Gln	Glm	Ala	Ala	Leu	Pro	Leu	Lev	Ser
				140					145		•			150
Asn	Ala	Glu	Cys	Lys	Lys	ser	urp	Gly	Arg	Ary	Ile	Thr	Asp	Val
			•	155			•		160					165
Mat	Ile	Cys	Ala	Gly	Ala	Ser	Gly	Val	Ser	Ser	Cys	Met	Gly	
				170	-			•	175					130
Ser	Gly	Gly	Pro	Leu	Val	Cys	Gln	Lys		Gly	Ala	Trp	Thŗ	
				185				<i>∴</i> .	190				·	195
Val	Gly	Ile	Val	Ser	Trp	Gly	Ser	ges.		Çye	Ser	Thr	Ser	
				200				•	205	,			1	210
Pro	Gly	Val	Tyr			Vai	Thr	Lys,		Ile	Pro	Trp	Awr	
		·		215	•	•			220				-	225
Lys	Ilë	Lev	Ala	Ala 230	Asti									
							-							
	<21			<b>7</b>										
		11>		255										-
		12>		PRT										
	421			Unkr	ÇŅM)									
	<22			<b>*</b> ********	<b>4</b>								. •	
		21>		DOMA					lask -	اگر بس	+1= + +	o. <del>F</del>	fart	Otho ji
	422	133								ಧ ವೆಲಾ ಕಾರ್ಡ				
•				(Fac	A) p	omo 1	OGOU	s to	eru	7' Y LT.	ODM	だずげ	وبل الاقتلب	ADG-15

	-:4	00>		7										
yığ	Ile	Val	Gly	•	Lys	Val	C}.≅	Pro		Gly	G1u	Cys	Pro	Trp
Gln	Val	Leu	Leu	5 Leu	Val	Asn	Gly	Ala	10 Gln	Leu	Суз	Gly	Cly	
	_			20			14 5		25	₩ 7 a'	175 -	e e e e e e e e e e e e e e e e e e e	Fals.s	30
Leu	Ile	Agn	Thr	11e	,lyth	Val	Vāi	ser	40 T3	WIG	urz	U Y M	rue	45
Lys	Ile	ГĀЗ	Asn	Trp	grīd	Asn	Ľéń	ll≘		Val	Læu	Gly	Glu	
_	<b>u</b>	.4	Glu	50	3.033	o <del>d</del> ler	<b>ገ</b> ኔ ሰቀታው	መኒክ	55 21 m	হুজুকু	<u> ጉ</u> ነተ	דוירב	val	50 Ala
qzA	Len	Ber	ATTE	65 65	rick.	GTŽ	Maju	GIG	70	٠.	137-3			75
Gln	Va1	Ile	Ile	Pro	Ser	Thr	Tyr	Val	Fro	Gly	Thr	Thr	Asn	
		٠		80					85			_	<i>4</i> 533	30 ·
Asp	Ile	Ala	Leu		Arg	L≘u	His	Gln	Pro 100	Val	Val	Γ≢π	Thr	Asp 105
Uin	57 1	W= 1	Pro	ůs Len	Cve	Ĩ, <b>≃</b> ¶1	Pro	( <del>†</del> ] 11		Thr	Phe	Ser	Glu	
TTD	121	й <del>С</del> т		110		20.7			115				· · ·	120
Thr	Leu	Als	Phe	Val	Arg	Fhe	Ser	Leu	Val	Ser	Gly	Trp	Gly	Gim .
				125					130					135
Leu	Leu	Asp	Arg	Gly	Ala	Thr	Ala			Leu	Met	Val	Leu	
				140					145	<b>4.7</b>	eta 9	ra	7. m.s.esa	150
Val	Fro	भाग्न	Leu	Met 155	Thr	Gln	ASD	Cys	160	GIN	GIN	adr	arg	165 ·
Mail	Glv	A rem	Ser		Alsm	Ile	Thr	Glu	-	Met	Ph∉	Çya	Ala	
y latent	*** *** ***	**************************************	<u></u> ,	170	<b>-</b>				175					180
Tyr	Ser	Asp	Gly	Ser	Lys	Asp	Sèr	Cys	Lys	Gly	Asp	Ser	Gly	Gly
				185					150					195
Fit	His	Alæ	Thr		$T_{\mathcal{Y}}r$	yrg	Gly	Thr		Tyr	Leu	Thr	Gly	
147	. <b>^</b> 4		-7.	200	· ·	C	<b>37</b> -	(Tiles ex	205	£2 las	uia	Tin e	ril v	210 vel
AST	ser	urp.	СlУ	215		广入党	12.T.C.	LIII	220	גדני	UTR.	1744	wary.	225
Tyrz	rdP	Arq	Val	•		Tyr	Ile	Glu		Leu	Ģ1n	Lys	Leu	•
-		_		230.		_			235	•				240
Yīğ	Ser	Glu	Pro	Arg 245	Fro	Gly	Val.	Leu	L∈u 250	Frg	Ala	Pro	Phe	Pro 255
	<23	10>		a						••				
•	<2.	11>		253									•	

	<2	12>		PRT					•					
	<2	13>		Unkr	OWI									
	-:2.	20>												
	· 42.	21>		IXMA	IN '									
	~ <b>2</b> ,	23>		Seri	ne p	rote	:35E	cata	dyti	ರ ಬೆಲ	main	of.	tiss	Пē
				plas	mino	gen	acti	vato	er (I	þa)	homó	logi	nis t	(D
				simi	lar	ෆ්ටගන	in i	II TA	IG-1	5				
	-< ₫	<00		8									•	
Aro	Ile	Lys	Gly	Gly	Leu	Phe	Ala	Asp	Ile	Ala	Ser	His	Pro	Trp
,		<b>_</b>	<b>-</b>	5					10					15
Gln	Ala	Ala	Il⊜	Pha	Ala	Lys	His	Arg	Ary	Ser	Pro	Gly	G1p	Arç
				20	•				25					30
Phe	Leu	Cys	Gly	Gly	Ile	Leu	I1e	Ser	Ser	్త్రొక	Trè	Ile	Leu	Ser
				35				-	40			٠		45
Ala	Als	His	Cys	Fhe	Gln	Glu	Arg	Phe	Pro	Pro	His	His	Leu	Thr
	•		-	50					55					80
Val	lie	Leu	Gly	Arg	Thr	Tyr	Arg	Val	Va1	Pro	Gly	Glu	Glu	Glu
		•	٠	65 -	•				70			-		75
Gln	Lys	Fhe	Glu	Val	Glu	Lys	Tyr	Ile	Val	His	Lys	Glu	Phe	Asp
				80					85		•			90
Asp	Asp	Thr	Tyr	Asp	Agn	Asp	Ile	2.1a	Fen	Leu	Gln	Leu	Lys	
	•			95					100					105
Asp	Ser	Ser	Mg	Cys	Ala	Gln	Glu	Ser	Ser	Val	Val	Arg	Thr	
				110					115			٠.		120
Cys	Leu	Fro	Fro	Alä	ysp	Lan	Gln	Leu		Asp	Trp	Thr	Glu	
				125		-			130					135
Glu	Leu	Sex	Gly		Gly	Lys	His	Glu		Leu	Ser	Pro	Fhe	
		٠		140	_		1	•	145	_	-	<b></b>	<b></b>	150
Ser	Glu	Arg	Γ≅α	•	Glu	Als	His	Val.		Leu	J. z.	5TQ	Ser	
				155			_		160	-1	**- 1	reels as	7	165
Ary	Çy₽	Thr	Ser	•	His	Leu	Lau	Asn			AFT	JIII	val	
		-	- 4	170	_	<b></b> 1	_	<b>~</b>	175.		F	(2.2 mg	27-	180
Met	Leu	Cys	ALa		asp	<u>11.17.</u>	wrg	ber		เจานั้	FIQ	ra TII	ute:	195
_	· · ·	_		185	رات رمشو		76	ه ــــــــــــــــــــــــــــــــــــ	190	- ـ ـ <u>-</u> -	D	Lass	57 <b>±</b> , 1	
Lan	His	Asp	Pla		GAT.	GIY	asp	SEL	205	مِيْلِيهِ ها	rrq		. 8. <del>22.</del> ⊤	210
				200			•		작다고					اها مار مند

### SUBSTITUTE SHEET (RULE 26)

```
Leu Asn Asp Gly Arg Met Thr Leu Val Gly Ile Ile Ser Trp Gly
                                                         225
                                     220
                215
Leu Gly Cys Gly Gln Lys Asp Wal Pro Gly Wal Tyr Thr Lys Wal
                                                         240
                                     235
                230
Thr Asn Tyr Leu Asp Trp Ile Arg Asp Asn Met Arg Pro
                                     250
                245
     <210>
                2900
     <211>
                MIF.
     <212>
                Homo sapiens
     <213>
     <220>
                SNC19 mWNA sequence (U20428)
     <223>
     <400>
cyctgggtgg tgctggtage egtgctgate ggcctcctct tgytcttgct
                                                          50
ggggatogge ttockggtyt ggcattbeca gtacegggae ytgeytgtec
                                                         100
                                                         150
agaaggtett caatggetac atgaggatea caaatgagaa ttttgtggat
                                                         200
gectacgaga actocazete caetgagttt gtaageetgg ceageaaggt
                                                         250
gaaggacgcg ctgaagctgc tgtacagegg agtoccattc ctgggcccct
                                                         300
accacaaqqa gtcggctgtg acggccttca gcgagggcag cgtcatcgcc
                                                         350
tactactggt otgagttcag catecogcag cacetggttg aggaggooga
                                                         400
gogogteato goosggagog ogtagteaty otgooccogo gogogogote
                                                         450
cotgaaqtoo titigigitoa cotcaqtigit qijotticoco acqigactoca
aaacagtaca gaggacccag gacaacagct gcagdtttgg ddtgdacgcc
                                                         500
goggtgtgga gotgatgege tteaccaege eggetteest gacagecet
                                                         550
                                                         £00
*cccccctca tgccccctgc cagtgggctg cggggacgcg acgcagtgct
                                                         $50
gagetacted agetgacted dagetegaet gegeetegae gagegeggea
                                                         700
gogacottggt gacgtgtaca acaccottgag coccatggag coccacgoot
                                                         750
ggfgagtgtg tggdaddtad cotocotoot acaaddigad cttddadtoo
                                                         800
etoccaegaa egteetgete ateacaetga taaccaacae tgacgeggea
                                                         850
teceggettt gaggesacet tetteraget geetaygatg ageagetgtg
gaggeegett acctaaagee caggggaeat teaacageee ctactaceca
                                                         900
ggccactacc cacceaacat tgactgcaca tggaaaattg aggtgcccaa
                                                         950
caaccagcat gtgaaqqtqc gcttcaaatt cttctacctg ctggagcccg 1000
gegtgeetge qqqeacetge cecaaggaet acqtqqaqat caatqqqqaq 1050
Aaatactgog gagagaggte coagttogto gtoaccagea acagoaacaa 1100
gatoacagtt cgcttccact cagatoagto ctacacogac accggottot 1150
tagotgaata cototootac gastocagtg accoatgood ggggsagtte 1200
acqtgccgca cggggcggtg tatccggaaq gagctgcgct gtgatggctg 1250
```

dacaactaca	cegaceacag	cgatgagete	aactgcagtt	gegaegeegg	1300
ccaccagtto			cangetette	tgggtctgcg	1350
acagogtyaa			acgascaque		1400
coggacccag		gitecaatgg		togaaasgoc	1450
	tääässääsc			<u>ರವಕ್ತವಿರದ್ದರ್ಭದ</u>	1500
_	tgaacytcyt			accgctgcct	1550
caatgggctc		<u> ಕಿದ್ದರ್ಭರಾವಕಿಂದರ</u>		යුගුසු <b>යෙකු</b> රුයේද	1600
	oggotoagat			gatgaggtea	1850
	aggetegtgt			atgagggdga	1700
	caggtaagcc			cacatotycy	1750
	catctctccc			acactgotac	1200
	gaggattcag			acggcettee	1850
		•		ಇದೆದೆಇಕೊಂಡಿಂತಕ	1900
	atcatctccc	accepticate	caatgacttc	accttcgact	1950
			cggcagagta	cagotocatu	2000
gtgeggddda	tetgeetgee	ggadgddtgd	catgtcttcc	ctgccggcaa	2050
ggocatotçq	gtcacqqqct	ggggacacac	ccegtatyga	ggcactggcg	2100
cgctgatcct	ೌರಶಾಶಾಥದಿದಿದ್ದ	gagateegeg	testessess	gaccacctge	2150
gagaaddtoc	£gccgcagca	gatesegeeg	cgcatgatgt	gegtgggött	2200
octoageqqe	ggcgtggact	catgaceggg	tgettccygg	ggacccctgt	2250
ರಂಜಭರಥ ಕಥಣ್ಣತ	ggcggatggg	cggatettee	aggooggtgt	66c6sactit	2300
ggagaegetg	cgctcagagg	ziaczagcca <u>u</u>	gegtgtacac	aaggeteett	2350
ctgtttcggg	aatggatcaa	agagaacact	gççqtatagg	<u> </u>	2400
acccaaatgt	gtacacctgc	ggggccaccc	atogtocaco	ccedtatace	2450
cgcctgcagg	ctggagactc	gogoscogty,	acctgcacca	<u>ತ್ರದಧಿದರ್ಭದಿಗೆ</u>	2500
acatacactg	tgaactcatc	tocaggoțoa	aatetgetag	aaaacctctc	2550
gottectesg	cctccaaagt	ggagetggga	gggtagaagg	ggaggaacac	2600
tggtggttat	actgacccaa	ctggggcaag	gtttgaagca	cagdbddggd	2650
agcccaagtg	ggcgaggacg	cgtttgtgca	tactgccctg	ctctstacac	2700
ggaagacctg	gatetetagt	gagtgtgact	googgatotg	getatautec	2750
ttggccacgc	ttsttgagga	agcccaggct	ದಡಿಡೆತಡೆದೆತ್ತುಂದ	tggaeaacag	2800
acomatetoa	mantuamant.	ootttaccaq	ctcccaggtg	acttcagtgt aaaaaaaaa	2850
<210>	10				
<211>	20				
<212>	DWA				
<213>		cial Sequen	ce		
<220>					

WO 99/42120 PCT/US99/03436

```
<221>
                primer
                Forward primer for analysis of overexpression
     <223>
                of TADG-15 menua by quantitative PCR.
     <4005
                10
stgacagagg attcaggtec
     <210>
                11
                20
     <211>
     <212>
                DNA
                Artificial Sequence
     <213>
     <220>
     <221>
                priner
                Reverse primer for analysis of overexpression
     <223>
                of TADG-15 mana by quantitative PCR.
     <400>
                11
                                  20
qaaqqtqaaq tdattqaaqa
     <210>
                12
     <211>
                17
     <212>
                LMA
                Artificial Sequence
     <213>
     <2205
                primer
     <221>
                Forward primer for analysis of $-tubulin mRNA.
     <2235
                empression by quantitative FCR.
     <400>
                12
                                 17
tgcattgara acgaggc
                13
     <210>
                17
     <211>
     <212>
                AM
                Artificial Sequence
     <213>
     <220>
                primer
     <221>
                Reverse primer for analysis of _-tubulin mPMR.
     <223>
                empression by quantitative PCR,
     <400>
                13
                                17
ctgtcttgac attgttg
```

										•				
	10 m	110>		14										
	· 42	11>		242										
		12>		FRT			-							
	مر کد په	13>		House	ı vəl	oien:	7			-				-
	∢	20>			-									
	<2	21>		DOME	AIM			•						
	*****	:23>		Seri	ine p	rote	:8 <b>36</b>	cetz	ılyti	ic di	mair	ು ಧರೆ	TALO	;-15,
	-: 4	<00		14	~									
Arg	Val	val	Gly	Gly	Thr	rep	Ala	Asp	Glu	Gly	Glu	Trp	Pro	Trp
		•		5					10					15
Gln	Val	Ser	Leu	His	Ala	Leu	Gly	Gln	Gly	His	Ile	Cys	Gly	Als
				20		٠.			25					30
Ser	Leu	Ile	Ser	Pro	Asn	Trp	Leu	Val	Ser	Ala	21a	His	Çys	Tyr
				35					4 Ç					45
Tie	Asp	Asp	Arg	Gly	Fhe	Arg	Tyr	Ser	Asp	Pro	Thr	Gln	Trp	
				50					55					\$Q
Ala	Phé	Leu	Gly	Leu	His	ASD.	Gln	Ser	Gln	Arg	Şer	Ala	Pro	
		•		65					70					75
Val	Gla	Glu	Arg	Arg	Leu	Lys	gra	Ile		Ser	His	Pro		
	-			80					85					.90 
Asn	¥\$Þ	Phe	Thr		Asp	Tyr	Asp	Il∈		Leu	Len	Ģlu	Len	
				95		•			100	_		_	_	105
Lys	Pro	Ala	Glu		\$er	Ser	Met	Val		Pro	I⊥≗	Cys	Leu	
			'	110	'				115	<b>.</b> 7	~1.	<b>-</b>	43m 7	120
Asp	Ala	Ser	His		Phe	FTO	YTS	Gly		WTS.	TTE	urp	TEV	
en ?		.e.a. 73	<b>5</b> - 1	125	<b>A</b> 1	<b></b>	ord i	m1	130	e33	70.7 m	7	<b>መ</b> ገመ	135
GLY	1 <del>, Li</del>	GTĀ	Hls		GLA	Jār	₽T <b>™</b>	G1y		ATA	Ala	TICIT	772	150
.44.7	<b>Y</b> 3 a.m.	<b></b> .		140	7	¥# 7	ትግ	* ana	145	пт ћ.,	CTÓ-10-1	J_4	Glu.	
13777	пЛя	ar.		155			TT=	Asn	160	7177	TILL	1-3-m	ملها ساد تتها	165
T mas	T man	There ex					T0 4 2-4.	Arg	,	Mat	( ^N t state	นะไ	ran se	
TIGIT	THEFT	BT O	GIII	170	772	7177	FLQ		175	Men	Oyu.	4 lekente	- u u y	180
T.=u	Ser	ദ്യ	ر کاری	_	Zer.	San	Over	Glr		Aan	Set	Glv	G1v	
and the top	ter that the	mark A.	Ψ¥.	185	1264	FOT	~g		190	water III.	W-5-	4+1		195
T.=11	Sar	()డిగా	Wall	•	<u> 21 =</u>	Aer.	GI w	Arg		Phe	ain	Ala	G1v	
حدد المناه	THE THE RE	<u></u>	₹ <b>Lb</b> ah	200			~ ± }	~ 'c1	205				<del>,</del>	210
ਰੋਜ਼ੀ	Ser	Time	Glw		Glv	Qu's	Ala	Çin	• •	Asn	Lvs	Pro	Glv	
v manaka	,	b	_	215	3	-y -			220		J <del></del> -		- <b>-</b> 4	225
									<u> </u>					•

# SEQ 16/17

### BUBSTITUTE SHEET (RULE 26)